CONTRIBUTIONS TO THE CALCULUS
OF VARIATIONS

THE UNIVERSITY OF CHICAGO PRESS
CHICAGO, ILLINOIS

THE BAKER & TAYLOR COMPANY
NEW YORK

THE CAMBRIDGE UNIVERSITY PRESS
LONDON

THE MARUZEN-KABUSHIKI-KAISHA
TOKYO, OSAKA, KYOTO, FUKUOKA, SENDAI

THE COMMERCIAL PRESS, LIMITED
SHANGHAI

CONTRIBUTIONS TO THE CALCULUS OF VARIATIONS

1931-1932

THESES SUBMITTED

TO

THE DEPARTMENT OF MATHEMATICS

OF THE

UNIVERSITY OF CHICAGO

THE UNIVERSITY OF CHICAGO PRESS ·
CHICAGO · ILLINOIS

PREFACE

The volume herewith presented contains theses concerning the calculus of variations submitted during the years 1931 and 1932 to the Department of Mathematics of the University of Chicago. It is a companion of the volume published two years ago containing theses in the same field submitted during 1930. The dissertations in both volumes were written under the direction of Professors G.A. Bliss and L.M. Graves. Of the papers here presented the one by Bliss and Hestenes, entitled "Sufficient conditions for a problem of Mayer in the calculus of variations," is not a dissertation. It has been included because of its intimate relationship with the immediately following thesis by Hestenes, and because its formulation was influenced in important particulars by ideas developed by Hestenes in his thesis work. The dissertation by Porter was submitted for the master's degree, and the others for the doctor's degree.

The preparation and publication of theses are topics of recurring interest in university circles upon which it is unlikely that complete agreement will ever be reached. To the writers of this preface it seems that the form of publication exemplified in this book has at the present time many advantages. It is a permanent but relatively inexpensive record, convenient for those

of us who have been most interested in the development of the subjects studied, as we hope it may be for others. If adopted widely it might relieve our journals of some of the congestion which has made their financial problems a source of continuous embarassment in the past.

The proper editing of a thesis for publication in a mathematical journal implies an intimacy of cooperation on the part of the faculty adviser which is likely to obscure the nature of the contribution made by the author himself and result in a paper written in the compressed mathematical style often adopted for journal publication. Neither of these results seems to us desirable. The first significant episode in the career of a scholar is the execution and publication of his first independently conceived research, to which he is impelled by his own studies and the inquiries in his own mind which they arouse. A doctoral dissertation is rarely such a publication, but it should be as close an approximation to one as possible. For this reason we feel that the candidate for the Ph. D. degree should be given as much freedom in his writing as is consistent with mathematical accuracy and clearness of style, even if the result is more verbose than could reasonably be approved by editors for publication in a journal.

When a number of students and faculty members are engaged in mathematical investigation it is inevitable that the subjects of researches begun at various times, and often delayed in execution, should show a considerable divergence of interest. An

explanation of the origins of the subjects of the various following theses would involve a historical sketch of the studies during several years past of the group interested in the calculus of variations at the University of Chicago. A remark may not be out of place, however, concerning four theses of a critical historical character which have been included, those of Sanger and the master's thesis by Porter in the present volume, and the theses of Huke and Duren in the volume which preceded. The reasons for our interest in these dissertations are numerous. All mathematical research should be accompanied if possible by a careful and extensive analysis of the literature. Such studies can only be made by specialists, and their results are frequently lost since there is no available medium of publication. We hope that over a period of years these and other similar theses may constitute an increasingly valuable contribution to the history of the calculus of variations. But most of all we value these analyses for their suggestiveness for further researches, and for the confidence which they give us in our orientations of present studies. We have had before us the example of Dickson's "History of the theory of numbers," whose influence in this respect has been remarkable.

The University of Chicago — G. A. BLISS

November, 1932. — L. M. GRAVES

TABLE OF CONTENTS

EDGE CONDITIONS FOR MULTIPLE INTEGRALS IN THE CALCULUS OF VARIATIONS

BY

JAMES ELLIS POWELL

TABLE OF CONTENTS

INTRODUCTION

Solutions of the simple integral problem of the calculus of variations are allowed for which y' has a finite number of ordinary discontinuities, called corners.* Weierstrass and Erdmann have shown that at a corner of a solution $y = y(x)$ the quantities $f_{y'}$ and $f - y'f_{y'}$ (where f is the integrand function) are continuous.** For the double integral problem in parametric form Kobb has obtained corresponding edge conditions.***

The purpose of this paper is to obtain conditions for the multiple integral problem analogous to the corner conditions for the simple integral problem or the edge conditions for the double integral problem. The integral to be minimized is

$$I = \int f(x,y,p)dx$$

where x is n-partite, y is m-partite, and

---

* L. M. Graves, Discontinuous solutions in the calculus of variations, Bulletin of the American Mathematical Society, XXXVI (1930), pp. 831-846.

** Weierstrass, Werke, VII, pp. 109, 250; Erdmann, Uber unstetige Losungen in der Variationsrechnung, Journal fur Mathematik, LXXXII (1876), pp. 21-30.

*** G. Kobb, Sur les maxima et les minima des integrales doubles, Acta Mathematica, (a) XVI (1892), p. 65, (b) XVII (1893), p. 321.

$$p = (p_{ij}) = \left(\frac{\partial y_j}{\partial x_i} \right).$$

Corresponding to an edge for the double integral problem we have an $(n-1)$-manifold, E', on which some of the partial derivatives p_{ij} are discontinuous. We derive $m+n$ conditions which the minimizing functions $y(x)$ must satisfy on E'.

In part I the statement of the problem of the calculus of variations for multiple integrals is set down, a fundamental lemma is proved, and the Lagrange differential equation is derived. In part II, the $m+n$ conditions, analogous to edge conditions, are derived. The argument is based on the fundamental lemma of part I. The method used in this part is one suggested by Professor G. A. Bliss. Parts III and IV deal similarly with the isoperimetric problem. In part V, the integrand function f is allowed to have finite discontinuities on a finite number of $(m+n-1)$-manifolds in the XY-space. The results of part II lead immediately to conditions on the minimizing functions $y(x)$ along the manifolds of intersection of $y(x)$ with the manifolds of discontinuity of f. Bliss and Mason*

* G. A. Bliss and M. Mason, A problem in the calculus of variations in which the integrand is discontinuous, Transactions of the American Mathematical Society, VII (1906), pp. 325-36.

treat this problem for the simple integral case. In parts VI and VII a second method is used to derive the edge conditions of part II. The method and argument used in these two parts were suggested to me by Professor L. M. Graves. In part VI a special case of the Mason fundamental lemma is proved for multiple integrals.* Part VII uses this lemma to derive the edge conditions.

---

* M. Mason, Beweis eines Lemmas der Variationsrechnung, Mathematische Annalen, LXI (1905), pp. 450-2.

I. THE LAGRANGE DIFFERENTIAL EQUATION

1. Notations; admissible manifolds. In this paper x will be n-partite and y will be m-partite. S will denote a bounded, closed, connected region of the X-space. Let M be the boundary of S. We shall denote by L a bounded (n-1)-manifold in XY-space whose projection on X is M. The partial derivatives $\left(\partial y_j / \partial x_i\right)$ will be denoted by p_{ij}. Here and in the following sections the index i will take the values 1 to n and the index j the values 1 to m. Let W be a region of the XYP-space. It will be convenient sometimes to let p denote the matrix of partial derivatives p_{ij}.

Except in parts VI and VII, a repeated index will denote a summation.

We first define a simple regular n-manifold $y = y(x)$, with boundary L, to be one such that, M, the projection of L, is cut by a line parallel to an axis in a finite number of points and segments, $y(x)$ is single valued and continuous in S, p is continuous in the interior of S, and the limit, over S, of p is continuous on M.

A regular n-manifold $y = y(x)$ is defined as one, such that $y(x)$ is continuous in S, and such that it consists of a finite number of cells which are simple regular

manifolds.

We take for an admissible manifold one which is regular and has its elements in W.

2. Statement of the problem of the calculus of variations for multiple integrals. Our problem of the calculus of variations is to find among all admissible manifolds

$$y = y(x), \qquad x \text{ in } S,$$

which pass through L, one which minimizes the n-tuple integral

$$I = \int_S f(x,y,p)\,dx \tag{1}$$

where, in W, f has as many continuous partial derivatives as desired.

3. Fundamental lemma. It is desirable and no more difficult to state and prove a lemma which is slightly more general than the strict extension of the one usually stated for the simple integral problem.

FUNDAMENTAL LEMMA. *If $N(x)$ is continuous in S except on a finite number of $(n-1)$-manifolds and*

$$\int_S N(x)\,\zeta(x)\,dx = 0$$

for all $\zeta(x)$ such that ζ is continuous in S, ζ_x

is continuous in S, *and* $\zeta = 0$ *on the boundary of* S, *then* $N(x) \equiv 0$ *in* S *except on the manifolds of discontinuity of* N. *If* $N(x)$ *has only ordinary discontinuities then* $N(x) \equiv 0$ *in* S.

We make the usual indirect proof. Suppose there exists an x_0 in S, not on one of the manifolds of discontinuity, such that $N(x) \neq 0$. Since there are only a finite number of manifolds of discontinuity it follows that $N(x)$ is continuous in the neighborhood of x_0. Therefore there exists a sphere Γ in S with radius ρ and center x_0 in which $N(x) \neq 0$. Let

$$\zeta(x) \equiv \begin{cases} [\rho^2 - \operatorname{mod}^2(x - x_0)]^2 & \text{in } \Gamma, \\ 0 & \text{outside } \Gamma, \end{cases}$$

where

$$\operatorname{mod}(x - x_0) \equiv \left(\sum (x_i - x_{i0})^2\right)^{\frac{1}{2}}.$$

This $\zeta(x)$ satisfies the three properties in the hypothesis of the lemma. However

$$\int_S N(x)\,\zeta(x)\,dx = \int_\Gamma N(x)\left[\rho^2 - \operatorname{mod}^2(x - x_0)\right]^2 dx \neq 0$$

which is a contradiction. It is evident that if $N(x)$ is allowed to have only ordinary discontinuities then $N(x) \equiv 0$ in S.

4. The Lagrange differential equation. Let

(2) $\mathfrak{M} \;:\; y = y(x), \quad x \text{ in } S,$

be a minimizing manifold for our problem.

THEOREM. On every part of $\mathfrak{M}$ where $y_k(x)$ has continuous second derivatives we have

(3) $$\frac{\partial}{\partial x_i} f_{p_{ik}} - f_{y_k} = 0.$$

For the proof let Γ be a sphere in S such that on the corresponding part of $\mathfrak{M}$, $y_k(x)$ has continuous second derivatives. Let $\zeta(x)$ be defined such that ζ is single valued and continuous in S, $\zeta = 0$ in $S - \Gamma$, and ζ_x is continuous in and on the boundary of Γ. The family of manifolds

(4) $$\left.\begin{aligned} y_j &= y_j(x), \quad j \neq k,\ j = 1, \dots, m, \\ y_k &= y_k(x) + a\,\zeta(x) \end{aligned}\right\} \equiv \bar{y}(x, a)$$

are admissible for a sufficiently small. Substituting (4) in (1) we get

$$I(a) = \int_S f(x, \bar{y}, \bar{p})\, dx,$$

and*

(5) $$I'(0) = \int_S \left(f_{y_k} \zeta + f_{p_{ik}} \zeta_{x_i} \right) dx = 0.$$

---

* Goursat-Hedrick, Mathematical Analysis, I, pp. 192-4

Making use of

$$\frac{\partial}{\partial x_i}\left(f_{p_{ik}}\zeta\right) = f_{p_{ik}}\zeta_{x_i} + \zeta\,\frac{\partial}{\partial x_i} f_{p_{ik}},$$

we have

$$\int_\Gamma f_{p_{ik}}\zeta_{x_i}\,dx = \int_\Gamma \frac{\partial}{\partial x_i}\left(f_{p_{ik}}\zeta\right)dx - \int_\Gamma \zeta\,\frac{\partial}{\partial x_i} f_{p_{ik}}\,dx$$

$$= -\int_\Gamma \zeta\,\frac{\partial}{\partial x_i} f_{p_{ik}}\,dx\,.$$

Therefore (5) reduces to

$$I'(0) = \int_\Gamma \zeta\left(f_{y_k} - \frac{\partial}{\partial x_i} f_{p_{ik}}\right)dx = 0\,.$$

Applying the fundamental lemma we obtain

$$f_{y_k} - \frac{\partial}{\partial x_i} f_{p_{ik}} = 0\,.$$

II. ANALOGUE OF THE EDGE CONDITIONS

5. Edges. Since the minimizing manifold (1) is admissible it consists of a finite number of cells in each of which the partial derivatives p exist and are continuous. Let E' be an $(n-1)$-manifold which is a common boundary of two of the cells. On E' some of the partial derivatives p_{ij} may be discontinuous. E' is then the analogue of a corner in the simple integral problem or of an edge in the double integral problem. In order that we can make use of equation (3) in the following work we require that $y(x)$ have continuous second derivatives on each side of E'. We also shall confine ourselves to that part of E' whose projection E on X can be represented by the equations

$$(6) \qquad x_i = \xi_i(u_1, \text{----}, u_{n-1}), \quad u \text{ in } U,$$

with ξ of class C''', and $\sum A_i^2 = 1$ where $A_i = (-1)^{i-1}$ times the determinant of the matrix obtained by deleting the i-th column of the matrix

$$\begin{pmatrix} \xi_{1u_1}, \text{-----------}, \xi_{nu_1} \\ \text{-------------} \\ \xi_{1u_{n-1}}, \text{---------}, \xi_{nu_{n-1}} \end{pmatrix}$$

We proceed to find necessary conditions on the minimizing functions (1) at points of E'. These conditions are the so-called "Edge Conditions".

6. <u>The normal coordinate system</u>. Introduce near E a uv-coordinate system determined by the equations

(7) $x_i = \xi_i(u) + vA_i$, u, v in B: u in U, $v_1 \leq v \leq v_2$

where v is one-partite and $v_1 < 0$ and $v_2 > 0$ are sufficiently small in absolute value that there exist unique functions

$$u_\ell = U_\ell(x_1, ----, x_n), \quad v = V(x_1, ----, x_n)$$

of class C'' satisfying equations (7). This is possible* since for x on E we have $(u_1, ---, u_{n-1}, v) = (u_1, ---, u_{n-1}, 0)$ and the functional determinant of (7)

$$\Delta = \begin{vmatrix} \xi_{1u_1} + vA_{1u_1}, & ---, & \xi_{nu_1} + vA_{nu_1} \\ ---- & ---- & ---- \\ \xi_{1u_{n-1}} + vA_{1u_{n-1}}, & ---, & \xi_{nu_{n-1}} + vA_{nu_{n-1}} \\ A_1 & ,---, & A_n \end{vmatrix}_{v=0} = \begin{vmatrix} \xi_{1u_1} & ,--, & \xi_{nu_1} \\ ---- & & ---- \\ \xi_{1u_{n-1}}, & --, & \xi_{nu_{n-1}} \\ A_1 & ,--, & A_n \end{vmatrix}$$

$$= \pm \sum A_i^2 = \pm 1 \neq 0.$$

---

* G. A. Bliss, Princeton Colloquium Lectures.

I have used B to denote a region in the UV-space and also the corresponding region in the X-space.

I have called this system normal because for $n = 2,3$, v measures along the normal.

Since v is negative on one side and positive on the other side of E, we shall, for convenience, take it as negative in C_1 and positive in C_2 where C_α $(\alpha = 1,2)$ are the two cells in S for which E is a part of the common boundary. Hereafter α will take only the values 1,2 and will never be used as a summation index.

7. <u>Family of variation manifolds</u>. In this section a family of variation manifolds is constructed corresponding to the edge E'.

Let

$$y = \psi(x), \quad x \text{ in } B, \tag{8}$$

with $\psi(x)$ of class C'', be an arbitrary manifold passing through E'. The transform of (8) by (7) is denoted by

$$y = \varphi(u,v), \quad u,v \text{ in } B. \tag{8'}$$

Denote the portions of the minimizing manifold $y(x)$ corresponding to the regions C_1 and C_2 by $y_1(x)$ and $y_2(x)$, respectively, and let their transforms by (7)

be denoted by

$$
(9) \qquad \begin{aligned} y &= Y_1(u,v), \quad u \text{ in } U, \quad v_1 \leq v \leq 0, \\ y &= Y_2(u,v), \quad u \text{ in } U, \quad 0 \leq v \leq v_2. \end{aligned}
$$

The notation y_α and Y_α are used here for the sets of functions $y_{j\alpha}$ and $Y_{j\alpha}$ respectively. We extend the range of definition of y_α and Y_α so that they are defined and of class C' in B by defining

$$
\begin{aligned} Y_1(u,v) &\equiv Y_1(u,0) + Y_{1v}(u,0)v, \quad u \text{ in } U, \; 0 < v \leq v_2, \\ Y_2(u,v) &\equiv Y_2(u,0) + Y_{2v}(u,0)v, \quad u \text{ in } U, \; v_1 \leq v < 0. \end{aligned}
$$

Next define an arbitrary one-parameter family of manifolds in B by the equation

$$
(10) \qquad v = a\,\zeta(u)
$$

where ζ is of class C' and $\zeta = 0$ on the boundary of U. The corresponding family of manifolds on

$$
y = \varphi(u,v)
$$

is given by the equations

$$
(11) \qquad \begin{aligned} v &= a\,\zeta(u), \\ y &= \varphi(u, a\,\zeta(u)). \end{aligned}
$$

Now construct on B two one-parameter families of

manifolds

$$y_{j\alpha} = \omega_{j\alpha}(x,a) \equiv$$

(12)

$$\frac{a\,\zeta(u)\,v_\alpha \int_0^1 [\varphi_{jv}(u,\theta v) - Y_{j\alpha v}(u,\theta v)]\,d\theta + [v_\alpha Y_{j\alpha}(u,v) - a\,\zeta(u)\,\varphi_j(u,v)]}{v_\alpha - a\,\zeta(u)},$$

$j = 1,\text{---},m$, $\alpha = 1,2$, u,v in B.

These families have the following properties: the manifolds (12) pass through the manifolds (11) for corresponding values of the parameter; the functions $\omega_{j\alpha}(x,a)$ reduce to $Y_{j\alpha}$ for u on the boundary of U or for $v = v_\alpha$; the functions $\omega_{j\alpha}(x,a)$ are continuous in B and at least of class C' in cells of B; for $a = 0$, $\omega_{j\alpha}(x,a)$ reduces to $Y_{j\alpha}$.

For the family of variation manifolds let

$$(13) \qquad \left.\begin{array}{lll} y = y(x) & \text{in} & S-B, \\ y = \omega_\alpha(x,a) & \text{in} & B_{\alpha a}, \end{array}\right\} \equiv \bar{y}_\alpha(x,a)$$

where $B_{\alpha a}$ is defined to be that part of B bounded by $v = v_\alpha$, the boundary of U, and $v = a\,\zeta(u)$. The functions defined by (13) are continuous in S, of class C' in cells, pass through the manifold L, and for corresponding values of the parameter a pass through the manifolds (11). For

$a=0$, (13) reduces to the minimizing manifold (2).

8. <u>The first variation</u>. Computing the value of I for the functions (13) we obtain

$$I(a) = H + \int_{B_{1a}} f(x,\bar{y}_1,\bar{p}_1)\,dx + \int_{B_{2a}} f(x,\bar{y}_2,\bar{p}_2)\,dx,$$

where H is independent of a and $B_{\alpha a}$ are defined above. It is noticed that the parameter a enters into both the integrand and the limits of integration. The method of finding $I'(0)$ in such a case is suggested in a paper of Simmons.* Separate the second and third integrals of $I(a)$ into two parts such that the first part has the parameter a only in the integrand. Next transform the second part by (7) into uv-coordinates.** This gives

$$I(a) = H + \int_{B_1} f\Big|^1 dx + \int_{B_2} f\Big|^2 dx$$
$$+ \int_U \int_0^{a\,\zeta(u)} f\Big|^1 |\Delta|\,dv\,du - \int_U \int_0^{a\,\zeta(u)} f\Big|^2 |\Delta|\,dv\,du,$$

---

*H. A. Simmons, First and second variations of a double integral for the case of variable limits, Transactions of the American Mathematical Society, XXVIII (1926), p. 235.

** J. Pierpont, Theory of Functions of a Real Variable, I, p. 552.

where Δ is defined on page 10, B_α is the part of C_α in B, and the superscripts 1,2 indicate the arguments of f. It is easy now to see that

(14) $$I'(0) = \int_{B_1} f_a\Big|^1 dx + \int_{B_2} f_a\Big|^2 dx$$

$$+ \int_U \int_0^0 f_a\Big|^1 |\Delta|\, dv\, du - \int_U \int_0^0 f_a\Big|^2 |\Delta|\, dv\, du$$

$$+ \int_U \xi(u) f\Big|^1 du - \int_U \xi(u) f\Big|^2 du$$

$$= \int_{B_1} f_a\Big|^1 dx + \int_{B_2} f_a\Big|^2 dx + \int_U \xi(u) f\Big|^1 du - \int_U \xi(u) f\Big|^2 du,$$

where

(15) $$f_a\Big|^\alpha = f_{y_j}\,\omega_{ja} + f_{p_{ij}}\,\omega_{jx_ia}\Big|^\alpha$$

Using

$$\frac{\partial}{\partial x_i}\left(f_{p_{ij}}\omega_{ja}\right) = f_{p_{ij}}\,\omega_{jx_ia} + \omega_{ja}\frac{\partial}{\partial x_i} f_{p_{ij}},$$

(14) reduces to

(16) $$I'(0) = \int_{B_1} \omega_{ja}\Big(f_{y_j} - \frac{\partial}{\partial x_i} f_{p_{ij}}\Big)\Big|^1 dx + \int_{B_2} \omega_{ja}\Big(f_{y_j} - \frac{\partial}{\partial x_i} f_{p_{ij}}\Big)\Big|^2 dx$$

$$+\int_{B_1} \frac{\partial}{\partial x_i}\left(f_{p_{ij}}\,\omega_{ja}\right)\Big|^1 dx \;+\int_{B_2} \frac{\partial}{\partial x_i}\left(f_{p_{ij}}\,\omega_{ja}\right)\Big|^2 dx$$

$$+\int_U \zeta(u)\, f\,\Big|^1 du \;-\int_U \zeta(u)\, f\,\Big|^2 du\,.$$

The first two integrals in (16) vanish by equation (3) since $y(x)$ is a minimizing function. We shall transform the third and fourth integrals.* The integral

$$\int_{B_1} \frac{\partial}{\partial x_i}\left(f_{p_{ij}}\,\omega_{ja}\right)\Big|^1 dx =$$

$$\sum_{i=1}^{n} \int\!\cdots\!\int \sum_{\sigma} \left\{\left[f_{p_{ij}}\,\omega_{ja}\right]^1_{x_i = z_{i2}^{(\sigma)}} - \left[f_{p_{ij}}\,\omega_{ja}\right]^1_{x_i = z_{i1}^{(\sigma)}}\right\} dx_1 \cdots dx_{i-1}\, dx_{i+1} \cdots dx_n$$

where $z_{i1}^{(\sigma)}$, $z_{i2}^{(\sigma)}$ are the values of x_i where a line parallel to the x_i-axis enters and leaves respectively the region B_1. By the definition of ω_j we notice that ω_{ja} is zero on the boundary of B_1 except for points on E. We need then to consider only the $z_{i1}^{(\sigma)}$, $z_{i2}^{(\sigma)}$ that are on E. We now transform by

(7) $\qquad x_i = \xi_i(u) + vA_i\,, \qquad u, v$ in B.

---

*E. Picard, Traite d'Analyse, 3-rd Ed., I, pp.165-6

We first show that A_1 is positive for $x_1 = z_{12}^{(\sigma)}$ and negative for $x_1 = z_{11}^{(\sigma)}$. We have supposed on page 11 that v is negative for points inside C_1 and therefore in B_1. Therefore v_{x_i} is positive for $x_1 = z_{12}^{(\sigma)}$ and negative for $x_1 = z_{11}^{(\sigma)}$ where we think of (7) as solved for u, v in terms of x. Next take the derivatives of (7) with respect to x_1 and set $v = 0$. This gives

$$0 = \xi_{1u_1} u_{1x_i} + - - - - + \xi_{1u_{n-1}} u_{n-1\,x_i} + v_{x_i} A_1 ,$$

$$- ,$$

$$1 = \xi_{iu_1} u_{1x_i} + - - - - + \xi_{iu_{n-1}} u_{n-1\,x_i} + v_{x_i} A_i ,$$

$$- ,$$

$$0 = \xi_{nu_1} u_{1x_i} + - - - - + \xi_{nu_{n-1}} u_{n-1\,x_i} + v_{x_i} A_n ,$$

where the left hand side is zero except for the i-th equation. Solving this set for v_{x_i} we get

$$v_{x_i} = \frac{\begin{vmatrix} \xi_{1u_1}, & \cdots, & \xi_{1u_{n-1}}, & 0 \\ \cdots, & \cdots, & \cdots, & \cdots \\ \xi_{iu_1}, & \cdots, & \xi_{iu_{n-1}}, & 1 \\ \cdots, & \cdots, & \cdots, & \cdots \\ \xi_{nu_1}, & \cdots, & \xi_{nu_{n-1}}, & 0 \end{vmatrix}}{\begin{vmatrix} \xi_{1u_1}, & \cdots, & \xi_{1u_{n-1}}, & A_1 \\ \cdots, & \cdots, & \cdots, & \cdots \\ \xi_{iu_1}, & \cdots, & \xi_{iu_{n-1}}, & A_i \\ \cdots, & \cdots, & \cdots, & \cdots \\ \xi_{nu_1}, & \cdots, & \xi_{nu_{n-1}}, & A_n \end{vmatrix}} = \frac{\pm A_i}{\pm \sum_{s=1}^{n} A_s^2} = A_i .$$

We see that in E, A_i has the same sign as v_{x_i}. This proves that A_i is positive at the points $x_i = z_{i2}^{(\sigma)}$ and negative at the points $x_i = z_{i1}^{(\sigma)}$. We have $A_i \neq 0$ except for those values of σ for which $z_{i1}^{(\sigma)} = z_{i2}^{(\sigma)}$. Transforming by (7) we therefore have*

$$\int_{B_1} \frac{\partial}{\partial x_i} \left(f_{p_{ij}}\, \omega_{ja}\right)\Big|' dx =$$

$$\sum_{i=1}^{n} \int \cdots \int \sum_{\sigma} \left\{ \left[f_{p_{ij}} \omega_{ja} \right]'_{x_i = z_{i2}^{(\sigma)}} - \left[f_{p_{ij}} \omega_{ja} \right]'_{x_i = z_{i1}^{(\sigma)}} \right\} dx_1 \cdots dx_{i-1} dx_{i+1} \cdots dx_n$$

---

* Cf. Pierpont, loc. cit.

$$= \sum_{i=1}^{n} \int \cdots \int \sum_{\sigma} \left\{ [f_{p_{ij}} \omega_{ja}]'_{x_i = z_{i2}^{(\sigma)}} |A_i| - [f_{p_{ij}} \omega_{ja}]'_{x_i = z_{i1}^{(\sigma)}} |A_i| \right\} du_1 \cdots du_{n-1}$$

$$= \sum_{i=1}^{n} \int \cdots \int \sum_{\sigma} \left\{ [f_{p_{ij}} \omega_{ja}]'_{x_i = z_{i2}^{(\sigma)}} A_i + [f_{p_{ij}} \omega_{ja}]'_{x_i = z_{i1}^{(\sigma)}} A_i \right\} du_1 \cdots du_{n-1}$$

$$= \int_U \omega_{ja} (f_{p_{ij}} A_i) \Big|^1 du .$$

Similarly

$$\int_{B_2} \frac{\partial}{\partial x_i} (f_{p_{ij}} \omega_{ja}) \Big|^2 dx = - \int_U \omega_{ja} (f_{p_{ij}} A_i) \Big|^2 du ,$$

the negative sign being due to the fact that v is positive for points in B_2. Substituting these expressions in (16) we obtain

$$(17) \quad I'(0) = \int_U \omega_{ja} (f_{p_{ij}} A_i) \Big|^1 du - \int_U \omega_{ja} (f_{p_{ij}} A_i) \Big|^2 du$$

$$+ \int_U \zeta(u) f \Big/^1 du - \int_U \zeta(u) f \Big/^2 du$$

$$= \int_U \left[\omega_{ja} (f_{p_{ij}} A_i) + \zeta(u) f \right]_2^1 du .$$

The value of ω_{ja} computed from equation (12) for $v=0$, $a=0$ is

$$\omega_{j\alpha a}\Big|_{\substack{a=0\\ v=0}} = \zeta(u)\left\{\varphi_{jv}(u,0) - y_{j\alpha v}(u,0)\right\} + \zeta(u)\left\{\frac{y_{j\alpha}(u,0)-\varphi_j(u,0)}{v_\alpha}\right\}.$$

Since $y=\varphi(u,v)$ passes through E', $\varphi_j(u,0) = Y_{j\alpha}(u,0)$. It is convenient to introduce $\psi_j(x)$ in place of $\varphi_j(u,v)$ and $p_{ij}A_i$ in place of $Y_{j\alpha v}$. This gives

$$\omega_{j\alpha a}\Big|_{\substack{a=0\\ v=0}} = \zeta(u)\left\{\psi_{jx_i}A_i - p_{ij}A_i\right\}.$$

Putting this value of ω_{ja} in (17) we have finally

(18) $$I'(0) = \int_U \zeta(u)\left[(\psi_{jx_i}A_i - p_{ij}A_i)(f_{p_{kj}}A_k) + f\right]_2' du.$$

9. <u>Edge conditions</u>. Since $y(x)$ is a minimizing function $I'(0)=0$ and we apply the fundamental lemma of page 5 to get

(19) $$\left[(\psi_{jx_i}A_i - p_{ij}A_i)(f_{p_{kj}}A_k) + f\right]_2' = 0 \quad \text{on } E.$$

The functions $\psi_j(x)$ are arbitrary but the quantities ψ_{jx_i} are not. Since $y=\psi(x)$ passes through E', we have

$$y_j(\xi(u)) = \psi_j(\xi(u)), \quad j=1,\dots,m.$$

Differentiating partially with respect to u_ℓ we get

(20) $\quad p_{ij}\, \xi_{i u_\ell} = \psi_{j x_i}\, \xi_{i u_\ell}\,, \quad \ell = 1,\dots,n-1, \quad j = 1,\dots,m.$

If for each j and each $k \neq \beta$ we multiply the equations of (20) by the cofactors of $\xi_{k u_\ell}$ in A_β and add, we obtain

$$p_{kj}\, A_\beta - p_{\beta j}\, A_k = \psi_{j x_k} A_\beta - \psi_{j x_\beta} A_k\,,$$

or, solved for $A_\beta\, \psi_{j x_k}$,

(21) $\quad A_\beta\, \psi_{j x_k} = p_{kj}\, A_\beta - p_{\beta j}\, A_k + \psi_{j x_\beta} A_k\,, \quad k = 1,\dots,\beta-1,\beta+1,\dots,n, \; j = 1,\dots,m.$

Evidently (21) holds also for $k = \beta$. Multiplying equation (19) by A_β and making use of (21) and $\sum A_i^2 = 1$, we get

(22) $\quad \psi_{j x_\beta} \left[f_{p_{kj}} A_k \right]_2^1 + \left[f A_\beta - p_{\beta j}\, f_{p_{kj}} A_k \right]_2^1 = 0\,,$

for $\beta = 1,\text{----},n$. Since $\psi_{j x_\beta}$, $j = 1,\text{----},m$ are arbitrary provided $A_\beta \neq 0$, we obtain the edge conditions

I $\quad f_{p_{kj}} A_k \Big|^1 = f_{p_{kj}} A_k \Big|^2, \quad j = 1,\dots,m,$

and hence, whether $A_\beta \neq 0$ or not,

II $\quad f A_\beta - p_{\beta j}\, f_{p_{kj}} A_k \Big|^1 = f A_\beta - p_{\beta j}\, f_{p_{kj}} A_k \Big|^2, \quad \beta = 1,\dots,n.$

It should be noticed that these $m+n$ edge conditions are not independent. From (21) we have

$$\text{(a)} \qquad p_{ij} A_\beta - p_{\beta j} A_i \Big|^1 = p_{ij} A_\beta - p_{\beta j} A_i \Big|^2$$

whether the manifold $\mathfrak{M}$ is a minimizing manifold or not. By (19) and (20) we can write

$$\text{(b)} \qquad f - A_i p_{ij} f_{p_{kj}} A_k \Big|^1 = f - A_i p_{ij} f_{p_{kj}} A_k \Big|^2 .$$

Using (a), (b) and the m edge conditions I, we obtain the n conditions II. Therefore there are only $m+1$ independent edge conditions. Although Kobb* obtained three edge conditions in the double integral problem, only two of them are independent.

10. Summary. Let $y = y(x)$ be a minimizing manifold for the n-tuple integral

$$I = \int_S f(x,y,p)dx, \quad p = (p_{ij}) = \left(\frac{\partial y_j}{\partial x_i}\right).$$

We suppose that S can be divided into cells in each of which the partial derivatives p exist and are continuous. Let E' be an $(n-1)$-manifold which is a part of the common boundary between two of the cells C_1 and C_2. On E' then some of the partial derivatives p_{ij} may be discon-

---

* Loc. cit.

tinuous. We assume that on each side of E', $y(x)$ has continuous second derivatives and that E, the projection of E' on X, can be represented by the equations

$$x_i = \xi_i(u_1, \cdots, u_{n-1}), \quad u \text{ in } U,$$

with ξ of class C''' and $\sum A_i^2 = 1$ where $A_i = (-1)^{i-1}$ times the determinant of the matrix obtained by deleting the i-th column from the matrix

$$\begin{pmatrix} \xi_{1u_1}, & \cdots, & \xi_{nu_1} \\ \cdots, & \cdots, & \cdots \\ \cdots, & \cdots, & \cdots \\ \xi_{1u_{n-1}}, & \cdots, & \xi_{nu_{n-1}} \end{pmatrix}$$

We have proved that at points on E' the following $m+n$ necessary conditions on the functions y_j must hold.

I $$f_{p_{ij}} A_i \Big|^1 = f_{p_{ij}} A_i \Big|^2, \quad j = 1, \cdots, m,$$

II $$f A_\beta - p_{\beta j} f_{p_{ij}} A_i \Big|^1 = f A_\beta - p_{\beta j} f_{p_{ij}} A_i \Big|^2, \quad \beta = 1, \cdots, n.$$

III. ISOPERIMETRIC PROBLEM; THE LAGRANGE MULTIPLIER RULE

The method used in the second part of this paper can be modified slightly so as to apply to the isoperimetric problem. We shall keep the same notations that we have used before.

11. Statement of the isoperimetric problem. The isoperimetric problem for multiple integrals is to find among all admissible manifolds

$$y = y(x), \quad x \text{ in } S,$$

which pass through L and which make

$$J = \int_S g(x,y,p)dx = j_0 \quad ,$$

where j_0 is a constant, one which minimizes the n-tuple integral

$$I = \int_S f(x,y,p)dx.$$

It is assumed that f and g both have, in W, as many continuous derivatives as desired.

12. The Lagrange multiplier rule. The method that Bliss used in the isoperimetric problem for simple

integrals is used here in the study of the multiple integral problem.*

Let

(23) $\mathfrak{M}$: $y = y(x)$, x in S,

be a minimizing manifold for our problem.

We first prove the following lemma.

LEMMA. *The determinant*

$$(24) \quad \begin{vmatrix} \int_S (f_{y_j} s_j + f_{p_{ij}} s_{j x_i}) dx, & \int_S (f_{y_j} t_j + f_{p_{ij}} t_{j x_i}) dx \\ \int_S (g_{y_j} s_j + g_{p_{ij}} s_{j x_i}) dx, & \int_S (g_{y_j} t_j + g_{p_{ij}} t_{j x_i}) dx \end{vmatrix}$$

must vanish for every set of functions $s(x)$, $t(x)$ *that are single valued and continuous in* S, *vanish on the boundary of* S, *and are of class* C' *in cells of* S.

The proof is the one that Bliss* gave for simple integrals. Suppose $s(x)$, $t(x)$ are functions for which the determinant is different from zero. We form the family of manifolds

(25) $y_j = y_j(x) + as_j(x) + bt_j(x) \equiv \bar{y}_j(x,a,b)$, $j = 1, ---, m$,

* Bliss, Lectures.

which are admissible for a,b sufficiently small, and consider the equations

$$
(26) \qquad
\begin{aligned}
I(a,b) &= \int_S f(x,\bar{y},\bar{p})dx = I(\mathfrak{M}) + z \,, \\
J(a,b) &= \int_S g(x,y,p)dx = j_0 \,,
\end{aligned}
$$

where $I(\mathfrak{M})$ is the value of I for the minimizing manifold (23). Equations (26) have the solution $(a,b,z) = (0,0,0)$ at which the value of the determinant

$$
\begin{vmatrix} I_a \,, & I_b \\ J_a \,, & J_b \end{vmatrix}
$$

is the determinant of the lemma which we have assumed different from zero. Therefore the equations (26) have the solution $a = a(z)$, $b = b(z)$ of class C' such that $a(0) = 0$, $b(0) = 0$. This means that for $z < 0$ the functions $a(z)$, $b(z)$ give manifolds of the family (25) for which $J = j_0$ and $I < I(\mathfrak{M})$. Since $\mathfrak{M}$ is a minimizing manifold the lemma is true.

By the use of this lemma we derive a Lagrange multiplier rule.

<u>THEOREM</u> I. <u>There exist constants</u> $(\lambda_1, \lambda_2) \neq (0,0)$ <u>such that on every part of</u> $\mathfrak{M}$ <u>where</u> $y_k(x)$ <u>has continuous second derivatives we have</u>

(27) $$\frac{\partial}{\partial x_i} h_{p_{ik}} - h_{y_k} = 0 \quad , k = 1, \ldots, m,$$

where $h = \lambda_1 f + \lambda_2 g$

If both elements of the second column of the determinant (24) are zero for every $t(x)$ then certainly both elements are zero for every $t(x)$ with $t_j(x) \equiv 0$, $j \neq k$, $j = 1, \ldots, m$. Then by sections 3 and 4 of part II, (27) is true for every pair (λ_1, λ_2). Consider then the case where there exists a $t(x)$ for which not both of these elements are zero. Select $(\lambda_1, \lambda_2) \neq (0,0)$ such that

$$\lambda_1 \int_S (f_{y_j} t_j + f_{p_{ij}} t_{jx_i}) dx + \lambda_2 \int_S (g_{y_j} t_j + g_{p_{ij}} t_{jx_i}) dx = 0 .$$

Since the determinant (24) vanishes we have

(28) $$\lambda_1 \int_S (f_{y_j} s_j + f_{p_{ij}} s_{jx_i}) dx + \lambda_2 \int_S (g_{y_j} s_j + g_{p_{ij}} s_{jx_i}) dx = \int_S (h_{y_j} s_j + h_{p_{ij}} s_{jx_i}) dx = 0$$

for every $s(x)$. Since (28) is true for every $s(x)$, it is true for every $s(x)$ with $s_j(x) \equiv 0$, $j \neq k$, $j = 1, \ldots, m$. Therefore (27) is true.

<u>THEOREM</u> II. <u>If the integral</u>

$$\int_S (g_{y_j} t_j + g_{p_{ij}} t_{jx_i})\,dx$$

<u>does not vanish for every</u> $t(x)$, <u>then</u> $\lambda_1 \neq 0$ <u>and there is but one set of the form</u> $(\lambda_1, \lambda_2) = (1, \lambda)$ <u>satisfying equation</u> (28).

It is evident from equation (28) that $\lambda_1 \neq 0$. If there exist two sets $(1, \lambda^{(1)})$, $(1, \lambda^{(2)})$ satisfying equation (28) then

$$(\lambda^{(1)} - \lambda^{(2)}) \int_S (g_{y_j} s_j + g_{p_{ij}} s_{jx_i})\,dx$$

for every $s(x)$, which is impossible by the hypothesis of the theorem.

IV. ANALOGUE OF THE EDGE COONDITIONS FOR THE ISOPERIMETRIC PROBLEM

13. Edges. As in part II let E' be the $(n-1)$-manifold which corresponds to an edge in the double integral problem. The projection E of E' on X is given by the equations

$$x_i = \xi_i(u), \quad u \text{ in } U,$$

with ξ of class C''' and $\sum A_i^2 = 1$ where A_i is defined in section 5. Let

$$y_j = \psi_j(x), \quad j = 1, \text{---}, m,$$

with $\psi(x)$ of class C'', be an arbitrary manifold through E'. Its transform by

$$(7) \qquad x_i = \xi_i(u) + v A_i, \quad u, v \text{ in } B,$$

is denoted by

$$y = \varphi(u, v), \quad u, v \text{ in } B.$$

14. Variation manifolds. In this section we pro-

ceed to construct a family of variation manifolds which are admissible.

Let

$$(29) \qquad v = a\,\zeta(u),$$

with ζ of class C' and $\zeta = 0$ on the boundary of U, be an arbitrary one-parameter family of manifolds in B. There is a corresponding family on $y = \varphi(u,v)$ given by

$$(30) \qquad \begin{aligned} v &= a\,\zeta(u) \\ y &= \varphi(u, a\,\zeta(u)). \end{aligned}$$

As in part II, we define functions

$$\omega_{j\alpha}(x,a) \equiv$$

$$(31) \qquad \frac{a\,\zeta(u)\,v_\alpha \int_0^1 [\varphi_{j v}(u,\theta v) - y_{j\alpha v}(u,\theta v)]\,d\theta + v_\alpha\, y_{j\alpha}(u,v) - a\,\zeta(u)\,\varphi_j(u,v)}{v_\alpha - a\,\zeta(u)}$$

We get a two-parameter family of admissible manifolds by letting

$$(32) \qquad \left.\begin{aligned} y &= y(x) + b\ t(x), \quad \text{in } S-B, \\ y &= \omega_\alpha(x,a) + b\ t(x), \quad \text{in } B_{\alpha a} \end{aligned}\right\} \equiv \bar{y}(x,a,b),$$

where $B_{\alpha a}$ is defined to be the part of B bounded by

$v = v_\alpha$, the boundary of U, and $v = a\,\zeta(u)$, and $t(x)$ is single valued and continuous in S, vanishes on the boundary of S, and is of class C' in cells of S. The manifolds (32) are continuous in S, of class C' in cells, pass through L, and for $b = 0$ and corresponding values of the parameter a pass through the manifolds (30). For $a = 0$, $b = 0$, (32) reduces to $y = y(x)$ in S.

15. __Variations of__ I __and__ J. Consider the two equations

$$I(a,b) = \int_{S-B} f(x,\bar{y},\bar{p})\,dx + \int_{B_{1a}} f(x,\bar{y}_1,\bar{p}_1)\,dx$$

$$+ \int_{B_{2a}} f(x,\bar{y}_2,\bar{p}_2)\,dx = I(\mathfrak{M}) + z\,,$$

$$J(a,b) = \int_{S-B} g(x,\bar{y},\bar{p})\,dx + \int_{B_{1a}} g(x,\bar{y}_1,\bar{p}_1)\,dx$$

$$+ \int_{B_{2a}} g(x,\bar{y}_2,\bar{p}_2)\,dx = j_0\,.$$

By the methods of part II, we find

$$I_a(0,0) = \int_{B_1} (f_{y_j}\,\omega_{ja} + f_{p_{ij}}\,\omega_{jax_i})|^1\,dx$$

(33)

$$+ \int_{B_2} (f_{y_j}\,\omega_{ja} + f_{p_{ij}}\,\omega_{jax_i})|^2\,dx + \int_U \zeta(u)\,f\Big|_2^1\,du\,,$$

and a similar expression for $J_a(0,0)$. The values of I_b and J_b are given by

$$I_b(0,0) = \int_S (f_{y_i} t_i + f_{p_{ij}} t_{ix_i})\, dx \,,$$

$$J_b(0,0) = \int_S (g_{y_i} t_i + g_{p_{ij}} t_{jx_i})\, dx \,.$$

The determinant

$$(34) \qquad \left| \begin{matrix} I_a \,, & I_b \\ J_a \,, & J_b \end{matrix} \right|_{\substack{a=0\\ b=0}}$$

must vanish for every $\zeta_i(u)$, $t(x)$ with ζ of class C' and $\zeta = 0$ on the boundary of U, and $t(x)$ continuous in S, $t = 0$ on the boundary of S, and of class C' in cells of S. The proof is the same as in section 12.

We shall, at first, assume that not both elements of the second column are zero for every $t(x)$. Therefore

$$\lambda_1 I_b + \lambda_2 J_b = 0 \,,$$

where the λ_1, λ_2 are defined on page 27. Since the determinant (34) vanishes we have

$$(35a) \qquad \lambda_1 I_a + \lambda_2 J_a = 0$$

for every $\zeta(u)$. If both elements of the second column

of (34) are zero for every $t(x)$, let

$$t(x) \equiv \omega_{j\alpha a}(x,0) \quad \text{in} \quad B_\alpha$$
$$\equiv 0 \quad \text{in} \quad S - B.$$

This $t(x)$ has the properties specified on page 25. Therefore

$$\lambda_1 I_b + \lambda_2 J_b = \lambda_1 I_a + \lambda_2 J_a = 0 . \tag{35b}$$

From (33), (35a), and (35b) we get

$$\int_{B_1} (h_{y_j}\, \omega_{ja} + h_{p_{ij}}\, \omega_{jax_i})\Big|^1 dx + \int_{B_2} (h_{y_j}\, \omega_{ja} + h_{p_{ij}}\, \omega_{jax_i})\Big|^2 dx + \int_U \zeta(u)\, h\Big|_2^1 du = 0 \tag{36}$$

for every $\zeta(u)$, where $h = \lambda_1 f + \lambda_2 g$.

We now make use of

$$\frac{\partial}{\partial x_i}\left(h_{p_{ij}}\, \omega_{ja}\right) = h_{p_{ij}}\, \omega_{jax_i} + \omega_{ja} \frac{\partial}{\partial x_i} h_{p_{ij}}$$

to change (36) into the form

$$\int_{B_1} \omega_{ja}\left(h_{y_j} - \frac{\partial}{\partial x_i} h_{p_{ij}}\right)\Big|^1 dx + \int_{B_2} \omega_{ja}\left(h_{y_j} - \frac{\partial}{\partial x_i} h_{p_{ij}}\right)\Big|^2 dx \tag{37}$$

$$+ \int_{B_1} \frac{\partial}{\partial x_i}\left(h_{p_{ij}}\, \omega_{ja}\right)\Big|^1 dx + \int_{B_2} \frac{\partial}{\partial x_i}\left(h_{p_{ij}}\, \omega_{ja}\right)\Big|^2 dx +$$

$$\int_U \zeta(u)\, h \Big|^1 du - \int_U \zeta(u)\, h \Big|^2 du = 0.$$

16. <u>Analogue of the edge conditions</u>. We notice that the left hand side of equation (37) is the same as the expression for $I'(0)$ in equation (16) except that f is replaced by $h = \lambda_1 f + \lambda_2 g$. The derivation of the edge conditions from (37) is then the same as that in part II. We have therefore the $m+n$ edge conditions:

$$\text{I} \qquad h_{p_{ij}} A_i \Big|^1 = h_{p_{ij}} A_i \Big|^2, \quad j = 1, \ldots, m,$$

$$\text{II} \qquad h A_\beta - p_{\beta j} h_{p_{ij}} A_i \Big|^1 = h A_\beta - p_{\beta j} h_{p_{ij}} A_i \Big|^2, \quad \beta = 1, \ldots, n.$$

V. CONDITIONS ON THE MINIMIZING FUNCTIONS ALONG A MANIFOLD OF DISCONTINUITY OF THE INTEGRAND FUNCTION

17. <u>Introduction.</u> In the study of the calculus of variations problem of minimizing the integral

$$\int_S f(x,y,p)dx \ , \quad x=(x_1,\text{--},x_n), \ y=(y_1,\text{--},y_m), \ p=(p_{ij}),$$

it is usually assumed that the integrand function f is continuous in its $m+n+mn$ arguments. In this part of the paper we shall allow the integrand function to have finite discontinuities on a finite number of $(m+n-1)$-manifolds in the XY-space. Conditions on the minimizing manifold

(38) $\quad \mathfrak{M} \ : \quad y=y(x) \ , \quad x \text{ in } S,$

are determined at the (n-1)-manifolds of intersection of (38) with the manifolds of discontinuity.*

---

* Simple integrals with discontinuous integrands have been studied by: G. A. Bliss and M. Mason, A problem in the calculus of variations in which the integrand is discontinuous, Transactions of the American Mathematical Society, VII (1906), pp. 325-36; E. J. Miles, Some properties of space curves minimizing a definite integral with discontinuous integrand, Bulletin of the American Mathematical Society, XX (1913), pp. 11-19; C. F. Roos, A general

18. __Statement of the problem.__ Let Q denote a bounded closed connected region of the XY-space whose projection on X is S. Let

$$(39) \qquad E'': \quad y_j = \tilde{y}_j(x_1,\text{--},x_n,t_1,\text{--},t_{m-1}), \quad j=1,\text{--},m ,$$

with y of class C'', be an $(m+n-1)$-manifold along which the integrand function may have finite discontinuities in its $m+n+mn$ arguments x,y,p. The manifold E'' separates Q into two regions Q_1, Q_2. We assume that in Q_1 and Q_2 the integrand function is continuous and has as many continuous partial derivatives as desired.

The problem of the calculus of variations here is to find among all admissible manifolds

$$y = y(x) \ , \quad x \ \text{in} \ S,$$

which pass through L and which cut E'' in a finite number of $(n-1)$-manifolds, the one that minimizes the sum of the two integrals

---

problem of minimizing an integral with discontinuous integrand, Transactions of the American Mathematical Society, XXXI (1929), pp. 58-70; E. H. Clarke, On the minimum of a sum of a definite integral and a function of a point, Dissertation, Chicago, 1922.

$$\int_{Q_1} f(x,y,p)dx \quad , \qquad \int_{Q_2} f(x,y,p)dx \; .$$

Let E' be an $(n-1)$-manifold of intersection of E'' with the minimizing manifold (38). We assume, that on each side of E' , $y(x)$ has continuous second derivatives. We confine our discussion to that part of E' whose projection E on X can be represented by the equations

$$x_i = \xi_i(v_1,\text{--},u_{n-1}) \; , \quad u \text{ in } U,$$

with ξ of class C''' and $\sum A_i^2 = 1$ where A_i is defined in section 5.*

19. <u>Edge conditions</u>. Let

$$t_r = t_r(x) \; , \quad r = 1,\text{---},m-1 \; , \tag{40}$$

be functions of class C'', arbitrary except that the manifold

$$y = \bar{y}(x,t(x)) \equiv \psi(x) \tag{41}$$

passes through E'. A normal coordinate system is set up near E as was done in section 6, and a family of variation

---

* It is only for convenience of notation that we have considered only one manifold E''. The argument holds for any finite number since we vary $y(x)$ only in the neighborhood of E'.

manifolds is defined as in section 7. The argument of section 8 is the same. We have then equation (19) in the form

$$(42) \qquad (\psi_{j x_i} A_i - p_{ij} A_i)(f_{p_{kj}} A_k) + f \Big|_2^1 = 0.$$

That is, the minimizing functions $y(x)$ must satisfy (42) along E'.

Computing $\psi_{j x_i}$ from (40) and (41), we get

$$(43) \qquad \psi_{j x_i} = \breve{y}_{j x_i} + \breve{y}_{j t_r} t_{r x_i}, \quad j = 1, \cdots, m, \quad i = 1, \cdots, n.$$

Putting in the values of $\psi_{j x_i}$ from (43), (42) takes the form

$$(44) \qquad \left\{ (\breve{y}_{j x_i} + \breve{y}_{j t_r} t_{r x_i}) A_i - p_{ij} A_i \right\} \left\{ f_{p_{kj}} A_k \right\} + f \Big|_2^1 = 0.$$

If we write (44) in the form

$$(45) \qquad \left\{ (\breve{y}_{j t_r} t_{r x_i}) A_i - (p_{ij} - \breve{y}_{j x_i}) A_i \right\} \left\{ f_{p_{kj}} A_k \right\} + f \Big|_2^1 = 0$$

we have again an expression of the form (19) with $\psi_{j x_i}$ replaced by $(\breve{y}_{j t_r} t_{r x_i})$ and p_{ij} replaced by $(p_{ij} - \breve{y}_{j x_i})$.

Since (41) passes through E', we have

$$\breve{y}_j(\xi, t(\xi)) = y_j(\xi), \quad j = 1, \cdots, m.$$

Differentiating partially with respect to u_ℓ we obtain

$$\breve{y}_{jx_i}\xi_{iu_l} + \breve{y}_{jt_r}t_{rx_i}\xi_{iu_l} = p_{ij}\xi_{iu_l}, \quad l=1,\dots,n-1, \; j=1,\dots,m,$$

which can be written as

$$(\breve{y}_{jt_r}t_{rx_i})\xi_{iu_l} = (p_{ij} - \breve{y}_{jx_i})\xi_{iu_l}, \quad l=1,\dots,n-1, \; j=1,\dots,m, \tag{46}$$

and finally as

$$\begin{aligned} A_\beta(\breve{y}_{jt_r}t_{rx_i}) &= (p_{ij} - \breve{y}_{jx_i})A_\beta - (p_{\beta j} - \breve{y}_{jx_\beta})A_i \\ &+ \breve{y}_{jt_r}t_{rx_\beta}A_i, \quad i=1,\dots,n, \; j=1,\dots,m, \end{aligned} \tag{47}$$

corresponding to equation (21). By the argument on page 21 we then have

$$\begin{aligned} &(\breve{y}_{jt_r}t_{rx_\beta})(f_{p_{ij}}A_i) + \\ &fA_\beta - (p_{\beta j} - \breve{y}_{jx_\beta})(f_{p_{ij}}A_i)\Big|_2^1 = 0, \quad \beta=1,\dots,n \end{aligned} \tag{48}$$

corresponding to equation (22). We write (48) in the form

$$\begin{aligned} &t_{rx_\beta}\{\breve{y}_{jt_r}(f_{p_{ij}}A_i)\} + \\ &\{fA_\beta - (p_{\beta j} - \breve{y}_{jx_\beta})(f_{p_{ij}}A_i)\}\Big|_2^1 = 0, \quad \beta=1,\dots,n. \end{aligned} \tag{49}$$

Since the functions $t_{r x_\beta}$, $r = 1, \dots, m-1$, are arbitrary, we have

I $$\breve{y}_{j t_r} f_{p_{ij}} A_i \Big|^1 = \breve{y}_{j t_r} f_{p_{ij}} A_i \Big|^2, \quad r = 1, \dots, m-1 ,$$

and

II $$f A_\beta - (p_{\beta j} - \breve{y}_{j x_\beta}) f_{p_{ij}} A_i \Big|^1 = f A_\beta - (p_{\beta j} - \breve{y}_{j x_\beta}) f_{p_{ij}} A_i \Big|^2 , \quad \beta = 1, \dots, n .$$

These $m+n-1$ conditions are the conditions that our minimizing functions $\dot{y}(x)$ must satisfy on the edge E'.

VI. MASON FUNDAMENTAL LEMMA

In the study of the simple integral problem two forms of the fundamental lemma are used.* The fundamental lemma of part I corresponds to one of these forms. Mason has proved a lemma for double integrals corresponding to the other form.** In this part of the paper we extend a special case of the Mason lemma to multiple integrals. Our hypotheses are slightly less restrictive. The proof is essentially the one that Mason used. In part VII edge conditions are derived by the use of this lemma.

The argument used in parts VI and VII was suggested to me by Dr. L. M. Graves.

---

* A. Huke, An historical and critical study of the fundamental lemma of the calculus of variations, Dissertation, Chicago, 1930.

**M. Mason, Beweis eines Lemmas der Variationrechnung, Mathematische Annalen, LXI (1905), pp. 450-2. Cf. also T. Kubota, Beweis der fundamentalen Hilfassatze in der Variationsrechnung, Tohoku Mathematical Journal, IX (1916), p. 191; A. Haar, Uber die Variation der Doppelintegrale, Journal fur Mathematik, CXLIX (1919), p. 1; and J. Schauder, Uber die Umkehrung eines Satzes aus der Variationsrechnung, Acta Litterarum ac Scientiarum, Szeged, IV (1928), pp.38-50.

20. Mason fundamental lemma. We adopt the notations and the definition of admissible manifolds of section 1. We shall also use the following notations.

A function $q(x)$ is said to be of class D on S if q is single valued and continuous in the interior of cells S_k of S, and the limit, over S_k, of q is continuous over the boundaries of S_k, where the boundaries of S_k are cut by a line parallel to an axis in a finite number of points and segments.

A function $\zeta(x)$ is said to be of class Z on S if, ζ is single valued and continuous on S, the partial derivatives $\zeta_{x_{i_1}\cdots x_{i_\gamma}}$, $i_1, \cdots, i_\gamma$ distinct, $\gamma = 1, \cdots, n$, are of class D in $(x_{i_1}, x_{i_2}, \cdots, x_{i_\gamma})$-space where on the manifolds of discontinuity the derivatives may be one sided.

In parts VI and VII, summation will be denoted by a summation sign and not by a repeated index.

FUNDAMENTAL LEMMA. If $N(x)$ is of class D on

$$R: \quad a_i \leq x_i \leq b_i, \quad i = 1, \cdots, n,$$

and the integral

$$\int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} N \zeta_{x_1 \cdots x_n} \, dx_1 \cdots dx_n$$

vanishes for every $\zeta(x)$ of class Z such that $\zeta = 0$ on the boundary of R, then there exist n functions

$$\varphi_i(x_1, \text{------}, x_{i-1}, x_{i+1}, \text{------}, x_n)$$

having only finite discontinuities on a finite number of (n-2)-manifolds in the $(x_1, ---, x_{i-1}, x_{i+1}, --- x_n)$-space, such that

(50) $$N(x) = \sum_i^{1,n} \varphi_i \quad \text{on} \quad R.$$

Also if the $N(x)$ is of class Z then the φ_i are of class Z in the $(x_1, ---, x_{i-1}, x_{i+1}, ---, x_n)$-space.

We notice first that for a function $\varphi_i(x_1, ---, x_{i-1}, x_{i+1}, ---, x_n)$, the integral

$$\int_{a_1}^{b_1} -- \int_{a_n}^{b_n} \varphi_i \, \zeta_{x_1 -- x_n} \, dx_1 -- dx_n =$$

$$\int_{a_1}^{b_1} -- \int_{a_{i-1}}^{b_{i-1}} \int_{a_{i+1}}^{b_{i+1}} -- \int_{a_n}^{b_n} \varphi_i \, \zeta_{x_1 -- x_{i-1} x_{i+1} -- x_n} \Big]_{a_i}^{b_i} dx_1 -- dx_{i-1} dx_{i+1} -- dx_n = 0,$$

since $\zeta = 0$ on the boundary of R.

We define

$$\zeta(x) = \Theta(x) +$$

$$\sum_{\beta}^{1,n} (-1)^{\beta} \sum_{\tau_1 < \tau_2 < \cdots < \tau_\beta}^{1,n} \frac{(x_{\tau_1} - a_{\tau_1}) \cdots (x_{\tau_\beta} - a_{\tau_\beta})}{(b_{\tau_1} - a_{\tau_1}) \cdots (b_{\tau_\beta} - a_{\tau_\beta})} \Theta_{\tau_1 \cdots \tau_\beta} ,$$

where

$$\Theta(x) \equiv \int_{a_1}^{x_1} \cdots \int_{a_n}^{x_n} N \, dx_1 \cdots dx_n ,$$

and $\Theta_{\tau_1 \cdots \tau_\beta}$ indicates that the upper limits $x_{\tau_1}, \cdots, x_{\tau_\beta}$ have been replaced by $b_{\tau_1}, \cdots, b_{\tau_\beta}$, respectively. This $\zeta(x)$ satisfies the hypotheses of the lemma. We compute

$$\zeta_{x_1 \cdots x_n} = N +$$

$$\sum_{\beta}^{1,n} (-1)^{\beta} \sum_{\tau_1 < \tau_2 < \cdots < \tau_\beta}^{1,n} \frac{1}{(b_{\tau_1} - a_{\tau_1}) \cdots (b_{\tau_\beta} - a_{\tau_\beta})} \int_{a_{\tau_1}}^{b_{\tau_1}} \cdots \int_{a_{\tau_\beta}}^{b_{\tau_\beta}} N \, dx_{\tau_1} \cdots dx_{\tau_\beta} .$$

Define

$$\varphi_i(x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n) \equiv \frac{1}{b_i - a_i} \int_{a_i}^{b_i} N \, dx_i +$$

$$\sum_{\beta}^{1,n} \frac{1}{\beta+1} (-1)^{\beta} \sum_{\tau_1 < \tau_2 < \cdots < \tau_\beta}^{\neq i,\ 1,n} \frac{1}{(b_i - a_i)(b_{\tau_1} - a_{\tau_1}) \cdots (b_{\tau_\beta} - a_{\tau_\beta})} \int_{a_i}^{b_i} \int_{a_{\tau_1}}^{b_{\tau_1}} \cdots \int_{a_{\tau_\beta}}^{b_{\tau_\beta}} N \, dx_i \, dx_{\tau_1} \cdots dx_{\tau_\beta} .$$

We have then

$$\int_R N \zeta_{x_1 \cdots x_n} dx = \int_R (N - \sum \varphi_i) \zeta_{x_1 \cdots x_n} dx =$$

$$\int_R (N - \sum \varphi_i)^2 dx = 0.$$

Therefore

$$N(x) = \sum_{i=}^{1,n} \varphi_i \text{ in } R.$$

The properties of the φ_i, stated in the lemma, follow from their definition.

VII. EDGE CONDITIONS OBTAINED BY THE USE OF THE MASON LEMMA

21. Edge conditions I. Let

$$y = y(x)\ , \quad x \text{ in } S, \tag{51}$$

be a minimizing manifold for our problem.

As in section 5 we let E' be an edge of our manifold (51) and E its projection on X. We again require that $y(x)$ have continuous second derivatives on each side of E'. In the equations

$$x_i = \xi_i(u_1,---,u_{n-1}) \tag{6}$$

of E, instead of requiring ξ_i to be of class C''', we require only that ξ_i be of class D'. Let

$$R: \quad a_i \leq x_i \leq b_i$$

be a rectangular region enclosing a part of E. We obtain edge conditions on the $y(x)$ along the corresponding part of E'.

We define a family of variation manifolds by

$$
(52) \qquad \begin{aligned} y &= y(x) \quad \text{in} \quad S - R, \\ y_j &= y_j(x), \; j \neq k, \; \text{in} \quad R, \\ y_k &= y_k(x) + a\zeta(x) \; \text{in} \quad R, \end{aligned} \qquad \equiv \quad \bar{y}(x,a)
$$

where ζ is of class Z on R and $\zeta = 0$ on the boundary and outside of R.

Putting (52) in the integral

$$\int_S f(x,y,p)dx$$

we get

$$I(a) = \int_S f(x,\bar{y},\bar{p})dx$$

and

$$(53) \qquad I'(0) = \int_R (f_{y_k}\zeta + f_{p_{ik}}\zeta_{x_i})dx = 0.$$

For convenience we define the functions

$$M = \int_{a_1}^{x_1} \cdots \int_{a_n}^{x_n} f_{y_k}\, dx_1 \cdots dx_n\,,$$

$$P_i = \int_{a_1}^{x_1} \cdots \int_{a_{i-1}}^{x_{i-1}} \int_{a_{i+1}}^{x_{i+1}} \cdots \int_{a_n}^{x_n} f_{p_{ik}}\, dx_1 \cdots dx_{i-1}\, dx_{i+1} \cdots dx_n \,, \quad i=1,\cdots,n.$$

The function M is of class Z and P_i is of class Z in the $(x_1,--,x_{i-1},x_{i+1},--,x_n)$-space. Also P_i is of class Z in the $(x_1,---,x_n)$-space near the edge by the hypothesis of the continuity of the second derivatives of $y(x)$ near the edge. Using

$$(M_{x_1--x_{n-1}}\zeta)_{x_n} = M_{x_1--x_n}\zeta + M_{x_1--x_{n-1}}\zeta_{x_n}$$

we get

$$\int_{a_n}^{b_n} M_{x_1--x_n}\zeta\, dx_n = -\int_{a_n}^{b_n} M_{x_1--x_{n-1}}\zeta_{x_n}\, dx_n,$$

since $\zeta = 0$ on the boundary of R. Continuing in this manner we get in the next step

$$\int_{a_{n-1}}^{b_{n-1}}\int_{a_n}^{b_n} M_{x_1--x_n}\zeta\, dx_{n-1}dx_n = \int_{a_{n-1}}^{b_{n-1}}\int_{a_n}^{b_n} M_{x_1--x_{n-2}}\zeta_{x_n x_{n-1}}dx_n dx_{n-1},$$

and finally

$$\int_R M_{x_1--x_n}\zeta\, dx = (-1)^n\int_R M\zeta_{x_1--x_n}\, dx.$$

In a similar manner we obtain

$$\int_{a_1}^{b_1}\cdots\int_{a_{i-1}}^{b_{i-1}}\int_{a_{i+1}}^{b_{i+1}}\cdots\int_{a_n}^{b_n} P_{i\,x_1\cdots x_{i-1}x_{i+1}\cdots x_n}\,\zeta_{x_i}\,dx_1\cdots dx_{i-1}\,dx_{i+1}\cdots dx_n$$

$$= (-1)^{n-1}\int_{a_1}^{b_1}\cdots\int_{a_{i-1}}^{b_{i-1}}\int_{a_{i+1}}^{b_{i+1}}\cdots\int_{a_n}^{b_n} P_i\,\zeta_{x_1\cdots x_n}\,dx_1\cdots dx_{i-1}\,dx_{i+1}\cdots dx_n .$$

Therefore we can write (53) in the form

$$I'(0) = \int_R (-1)^n (M - P_1 - P_2 - \cdots - P_n)\,\zeta_{x_1\cdots x_n}\,dx = 0.$$

By the fundamental lemma of section 20 we have

(54) $$N = M - \sum_i^{1,n} P_i = \sum_i^{1,n} \varphi_i$$

where φ_i does not contain the variable x_i. It should be noticed that $N(x)$ is of class Z near the edge B.

Letting

(55) $$V_i(x) \equiv \frac{1}{n}\int_{a_i}^{x_i} f_{y_k}\,dx_i - f_{p_{ik}}, \quad i = 1, \cdots, n,$$

we get (54) in the form

$$\sum_i \int_{a_1}^{x_1}\cdots\int_{a_{i-1}}^{x_{i-1}}\int_{a_{i+1}}^{x_{i+1}}\cdots\int_{a_n}^{x_n} V_i\,dx_1\cdots dx_{i-1}\,dx_{i+1}\cdots dx_n = \sum_i^{1,n}\varphi_i .$$

Let functions Q_i be defined by

$$Q_1 \equiv \int_{a_2}^{x_2} \cdots \int_{a_n}^{x_n} V_1 \, dx_2 \cdots dx_n - \varphi_n ,$$

$$Q_i \equiv \int_{a_1}^{x_1} \cdots \int_{a_{i-1}}^{x_{i-1}} \int_{a_{i+1}}^{x_{i+1}} \cdots \int_{a_n}^{x_n} V_i \, dx_1 \cdots dx_{i-1} dx_{i+1} \cdots dx_n ,$$

$$Q_n \equiv \int_{a_1}^{x_1} \cdots \int_{a_{n-1}}^{x_{n-1}} V_n \, dx_1 \cdots dx_{n-1} - \sum_{i=1}^{n-1} \varphi_i .$$

We have then

$$(56) \qquad \sum_i^{1,n} Q_i \equiv 0 \qquad \text{on } R.$$

The partial derivatives

$$(57) \qquad \begin{aligned} Q_{1\,x_2 \cdots x_{n-1}} &= \int_{a_n}^{x_n} V_1 \, dx_n - \varphi_{n\,x_2 \cdots x_{n-1}} \equiv \bar{q}_1 , \\ Q_{i\,x_1 \cdots x_{i-1} x_{i+1} \cdots x_{n-1}} &= \int_{a_n}^{x_n} V_i \, dx_n \equiv \bar{q}_i , \quad i = 2, \cdots, n-1, \end{aligned}$$

are continuous across the edge E except at a point on a line parallel to the x_n-axis having a segment in common with E or some other part of the boundaries of the cells S_k of S. The rectangle R can be taken sufficiently small about the point O, of E, in consideration such that

a line parallel to any x_i-axis, which does not include the point O in a segment in commom with E, will not have a segment in common with E or any other part of the boundaries of S_k in R. If a line parallel to the x_n-axis has a segment in common with E then $A_n = 0$ along this segment, by the argument in section 8 of part II. Since $A_i \neq 0$ for some i we can modify the proof from here on for those points of E. We assume then that the $\bar{q}_i$ are continuous across the edge E. Therefore $\partial/\partial u_\ell(\bar{q}_i)$ has the same value on both sides of E for $i = 1,--,n-1$, $\ell = 1,--,n-1$. We have then, except at edges of E,

$$(58) \qquad \bar{q}_{ix_1}\xi_{1u_\ell} + --- + \bar{q}_{ix_{n-1}}\xi_{n-1\,u_\ell} + V_i\,\xi_{n\,u_\ell}\Big|_2^1 = 0$$

for $\ell = 1,--,n-1$. From (58) we obtain

$$(59) \qquad A_n\,\bar{q}_{ix_i}\Big|_2^1 = A_i\,V_i\Big|_2^1, \quad i = 1,--,n-1.$$

Since either

$$\int_{a_i}^{x_i} f_{y_k}\,dx_i$$

is continuous across the edge E or $A_i = 0$ we have

$$A_i\int_{a_i}^{x_i} f_{y_k}\,dx_i\Big|_2^1 = 0$$

and therefore, from (55), (57), and (59)

$$A_n Q_{i\, x_1 \cdots x_{n-1}} \Big|_2^1 = A_i f_{p_{ik}} \Big|_2^1 , \quad i = 1, \cdots, n-1.$$

Also

$$Q_{n\, x_1 \cdots x_{n-1}} \Big|_2^1 = f_{p_{nk}} \Big|_2^1 .$$

From (56)

$$A_n \sum_i Q_{i\, x_1 \cdots x_{n-1}} = 0 .$$

Therefore we have the m edge conditions

$$(60) \qquad \sum_i^{1,n} A_i f_{p_{ik}} \Big|^1 = \sum_i^{1,n} A_i f_{p_{ik}} \Big|^2 , \quad k = 1, \cdots, m .$$

Since the number of edges of E is finite we still have the result (60) at these points by the continuity of $\sum_i A_i f_{p_{ik}}$ in the cells S_k of S.

22. <u>Other edge conditions</u>. In this section we obtain other edge conditions by a change of our problem into parametric form. Let

$$x_i = x_i(t_1, \cdots, t_n) , \qquad i = 1, \cdots n.$$

Then

$$y(x) = y(x(t)) = \bar{y}(t) ,$$

and

$$\bar{y}_{kt_\beta} = \sum_i y_{kx_i} x_{it_\beta} \quad , \quad k = 1, --, m, \quad \beta = 1, --, n.$$

Suppose

$$\Delta = | x_{it_\beta} | \neq 0.$$

Then

$$p_{ik} = y_{kx_i} = \frac{\Lambda_{ki}}{\Delta}$$

where Λ_{ki} is the determinant obtained from Δ by replacing the i-th row by $\bar{y}_{kt_\beta}$. Our integral I then takes the form

$$I = \int_S f\left(x, y, \frac{\Lambda}{\Delta}\right) \Delta \, dt, \quad \Lambda \equiv (\Lambda_{ki}),$$

$$= \int_S F(t, x, y, q, r) \, dt,$$

where $q \equiv (x_{it_\beta})$ and $r \equiv (\bar{y}_{kt_\beta})$, which is of the same form as before except that y is (m + n)-partite. Therefore, by (60), we can write

$$\begin{aligned} &\sum_\beta^{1,n} F_{q_{i\beta}} \bar{A}_\beta \Big|^1 = \sum_\beta^{1,n} F_{q_{i\beta}} \bar{A}_\beta \Big|^2, \quad i = 1, --, n, \\ (61) \qquad &\sum_\beta^{1,n} F_{r_{k\beta}} \bar{A}_\beta \Big|^1 = \sum_\beta^{1,n} F_{r_{k\beta}} \bar{A}_\beta \Big|^2, \quad k = 1, --, m. \end{aligned}$$

The $\bar{A}_\beta$ is defined in the same manner as the A_i, where E is given by the equations

$$t_\beta = \bar{t}_\beta(u_1, ---, u_{n-1}), \qquad \beta = 1, --, n.$$

Now let $x_i = t_i$. Then

$$\Delta = 1, \quad \bar{A}_\beta = A_\beta, \quad \bar{y}_{k t_\beta} = y_{k x_\beta} .$$

The partial derivatives $F_{q_{i\beta}}$ become

$$F_{q_{i\beta}} = F_{x_{i t_\beta}} = f \frac{\partial \Delta}{\partial x_{i t_\beta}} + \Delta \sum_{\gamma k} f_{p_{\gamma k}} \frac{\partial \left(\frac{\Lambda_{k\gamma}}{\Delta} \right)}{\partial x_{i t_\beta}}$$

$$= f \frac{\partial \Delta}{\partial x_{i t_\beta}} + \frac{1}{\Delta} \sum_{\gamma k} f_{p_{\gamma k}} \left\{ \Delta \frac{\partial \Lambda_{k\gamma}}{\partial x_{i t_\beta}} - \Lambda_{k\gamma} \frac{\partial \Delta}{\partial x_{i t_\beta}} \right\}.$$

Putting in the values

$$\Delta = 1, \quad \frac{\partial \Delta}{\partial x_{i t_\beta}} = \delta_{i\beta}, \quad \frac{\partial \Lambda_{ki}}{\partial x_{i t_i}} = 0, \quad \frac{\partial \Lambda_{k\gamma}}{\partial x_{i t_i}} = p_{\gamma k}, \; \gamma \neq i,$$

$$\Lambda_{k\gamma} = p_{\gamma k}, \quad \frac{\partial \Lambda_{k\gamma}}{\partial x_{i t_\beta}} = - \delta_{\gamma\beta} \, p_{ik}, \quad \beta \neq i,$$

we get

$$F_{q_{ii}} = f - \sum_k p_{ik} f_{p_{ik}},$$

and

$$F_{q_{i\beta}} = -\sum_{k} p_{ik} f_{p_{\beta k}}, \quad (i \neq \beta).$$

Similarly

$$F_{r_{k\beta}} = f_{p_{\beta k}} .$$

Substituting these values in (61) we obtain the m+n edge conditions

$$\text{I} \qquad \sum_{i}^{1,n} f_{p_{ik}} A_i \Big|^1 = \sum_{i}^{1,n} f_{p_{ik}} A_i \Big|^2, \quad k = 1, \cdots, m$$

and

$$\text{II} \qquad f A_\beta - \sum_{k}^{1,m} p_{\beta k} \sum_{i}^{1,n} f_{p_{ik}} A_i \Big|^1 = f A_\beta - \sum_{k}^{1,m} p_{\beta k} \sum_{i}^{1,n} f_{p_{ik}} A_i \Big|^2, \quad \beta = 1, \cdots, n.$$

These are the conditions that we found in part II.

BIBLIOGRAPHY*

A. GENERAL

1. L. M. Graves, Discontinuous solutions in the calculus of variations, Bulletin of the American Mathematical Society, XXXVI (1930), pp. 831-46.

2. H. A. Simmons, First and second variations of a double integral for the case of variable limits, Transactions of the American Mathematical Society, XXVIII (1926), p. 235.

3. C. A. Fischer, Minima of double integrals with respect to one-sided variations, Annals of Mathematics, (2), XVI (1915), pp. 162-8.

B. CORNER AND EDGE CONDITIONS

4. Weierstrass, Werke, VII, pp. 109, 250.

5. Erdmann, Uber unstetige Losungen in der Variationsrechnung, Journal fur Mathematik, LXXXII (1876),pp.21-30

---

* For a complete bibliography on discontinuous solutions see paper (1). This paper gives a brief resume of the work that has been done on discontinuous solutions.

6. Caratheodory, Uber die diskontinuierlichen Losungen in der Variationsrechnung, Dissertation, Gottingen, 1904, (v + 71 pp.).

7. G. Kobb, Sur les maxima et les minima des integrales doubles, Acta Mathematica, (a) XVI (1892), p. 65, (b) XVII (1893), p. 321.

C. DISCONTINUOUS INTEGRANDS

8. Bliss and Mason, A problem in the calculus of variations in which the integrand is discontinuous, Transactions of the American Mathematical Society, VII (1906), pp. 325-36.

9. E. J. Miles, Some properties of space curves minimizing a definite integral with discontinuous integrand, Bulletin of the American Mathematical Society, XX (1913), pp. 11-19.

10. C. F. Roos, A general problem of minimizing an integral with discontinuous integrand, Transactions of the American Mathematical Society, XXXI (1929), pp. 58-70.

11. E. H. Clarke, On the minimum of the sum of a definite integral and a function of a point, Dissertation, Chicago, 1922.

D. MASON FUNDAMENTAL LEMMA

12. M. Mason, Beweis eines Lemmas der Variationsrechnung, Mathematische Annalen, LXI (1905), pp. 450-2.

13. T. Kubota, Beweis der fundamentalen Hilfassatze in der Variationsrechnung, Tohoku Mathematical Journal, IX (1916), p. 191.

14. A. Haar, Uber die Variation der Doppelintegrale, Journal fur Mathematik, CXLIX (1919), p. 1.

15. J. Schauder, Uber die Umkehrung eines Satzes aus der Variationsrechnung, Acta Litterarum ac Scientiarum, Szeged, IV (1928), pp. 38-50.

16. A. Huke, An historical and critical study of the fundamental lemma of the calculus of variations, Dissertation, Chicago, 1930.

THE EULER-LAGRANGE MULTIPLIER RULE
FOR DOUBLE INTEGRALS

BY

MAX CORAL

TABLE OF CONTENTS

INTRODUCTION

The following pages are devoted to a discussion of certain special cases of a problem of Lagrange in the calculus of variations for double integrals; more specifically, to the establishment of an analogue of the famous Euler-Lagrange multiplier rule. For simpler integrals, the problem of Lagrange is that of finding in a prescribed class of arcs $y_i = y_i(x)$ $(x_1 \leq x \leq x_2;\ i = 1,2,\ldots,n)$ which join the two fixed points (x_1,y_1), (x_2,y_2) and satisfy a system of differential equations $\phi_\alpha(x,y_1,\ldots,y_n,y_1',\ldots,y_n') = 0$ $(\alpha = 1,2,\ldots,m)$ one which minimizes the integral

$$I = \int_{x_1}^{x_2} f(x,y_1,\ldots,y_n,y_1',\ldots,y_n')\,dx.$$

According to the multiplier rule, discovered independently by Euler and Lagrange, there exist a constant λ_0 and m functions $\lambda_\alpha(x)$ such that along an arc which furnishes a solution for the above problem the equations $(d/dx)F_{y_i'} - F_{y_i} = 0$ hold for the function $F = \lambda_0 f + \lambda_\alpha \phi_\alpha$.

An analogous problem for double integrals would be: to minimize the integral

$$I = \iint_A f(x,y,z_1,\ldots,z_n,p_1,\ldots,p_n,q_1,\ldots,q_n)\,dxdy$$

in a class of surfaces $z_i = z_i(x,y)$ $[(x,y)$ in A; $i = 1,2,\ldots,n]$, all of which pass through the same closed curve above the

boundary of the region A and which satisfy the system of partial differential equations $\phi_\alpha(x,y,z_i,p_i,q_i) = 0$, m in number, where it is understood that $p_i = (\partial/\partial x) z_i$, $q_i = (\partial/\partial y) z_i$. It is characteristic of the incompleteness existing in the calculus of variations for double integrals that the multiplier rule for a problem of such generality has not yet been established. Apparently the only discussion in the literature of a problem of Lagrange for double integrals is that to be found in a paper by Gross[1], who treated the case of a surface in the space of (x,y,z_1,z_2)-points, defined over a square B of the (x,y)-plane, which minimizes the integral

$$I = \iint_B f(x,y,z_1,z_2,p_1,p_2,q_1,q_2)\,dxdy$$

in a class of surfaces $z_i = z_i(x,y)$ defined in B, having the same boundary as the minimizing surface, and satisfying the equation $\phi(x,y,z_1,z_2,p_1,p_2) = 0$. Gross was able, under certain hypotheses, to prove the existence of a function $\lambda(x,y)$ such that along the minimizing surface the Lagrange partial differential equations $F_{z_i} - (\partial/\partial x)F_{p_i} - (\partial/\partial y)F_{q_i} = 0$ hold, where $F = f + \lambda\phi$.

In the first chapter below, the case considered by Gross will be re-examined and a much simpler proof of the multiplier rule made for it under restrictions less severe than those imposed by him. The procedure will consist essentially in the

---

[1] W. Gross, Das isoperimetrische Problem bei Doppelintegrale, Monatshefte für Mathematik und Physik, XXVII(1916), pages 114-120.

derivation of an auxiliary minimum problem, a problem of Lagrange for simple integrals, which can be treated by known methods. A modification of this method will be found successful, in the second chapter, in the discussion of a somewhat more general problem, for which the equation of condition has the form $\phi(x,y,z_1,z_2,p_1,p_2,q_2) = 0$. However, a proof for the most general type of equation $\phi(x,y,z_1,z_2,p_1,p_2,q_1,q_2) = 0$, in which all four derivatives appear, is still lacking.

CHAPTER I

THE MULTIPLIER RULE FOR THE CASE STUDIED BY GROSS

1. Formulation of the problem. Consider a surface

$$z_i = Z_i(x,y) \qquad (i = 1,2) \tag{1.1}$$

defined in and on the boundary of a region B of the (x,y)-plane. In this chapter the multiplier rule will be proved for a surface (1.1) which furnishes a minimum for the integral

$$I = \iint_B f(x,y,z_1,z_2,p_1,p_2,q_1,q_2)\,dxdy \tag{1.2}$$

in a class of surfaces

$$z_i = z_i(x,y) \qquad [(x,y) \text{ in } B;\ i = 1,2] \tag{1.3}$$

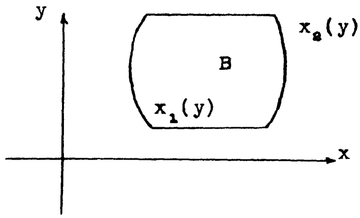
all of which pass through the same closed curve as does the surface (1.1) above the boundary of B and which satisfy the equation

$$\phi(x,y,z_1,z_2,p_1,p_2) = 0. \tag{1.4}$$

In the integrand of I and in the function ϕ it is understood that for p_i, q_i are to be substituted respectively the derivatives $z_{ix}(x,y)$, $z_{iy}(x,y)$.

The region B is to consist of all points (x,y) which satisfy the conditions: $y_1 \leqslant y \leqslant y_2$, $x_1(y) \leqslant x \leqslant x_2(y)$; the functions $x_i(y)$ $(i = 1,2)$ are continuous together with their

derivatives on the interval $y_1 y_2$ and have $x_1(y) < x_2(y)$ for every y of the interval.



We proceed under the following hypotheses:

(a) the functions $Z_i(x,y)$ defining the minimizing surface are of class C'' in B[1];

(b) there is a region R of elements (x,y,z_i,p_i,q_i) to which all the elements $[x,y,Z_i(x,y),Z_{ix}(x,y),Z_{iy}(x,y)]$ having (x,y) in B are interior and in which the functions f and ϕ are defined and of class C'' at least;

(c) at no point of the minimizing surface do both ϕ_{p_1} and ϕ_{p_2} vanish simultaneously.

It follows from (c) that there exists a function $\psi(x,y,z_1,z_2,p_1,p_2)$, defined and of class C'' in R, such that the determinant

$$(1.5) \qquad \begin{vmatrix} \phi_{p_1} & \phi_{p_2} \\ \psi_{p_1} & \psi_{p_2} \end{vmatrix}$$

---

[1] After Bolza, a function is said to be of class $C^{(n)}$ in its region of definition if its partial derivatives of all orders up to and including the nth exist and are continuous in the region.

is never zero along the minimizing surface. For example, the function $\psi^* = \phi_{p_1} p_2 - \phi_{p_2} p_1$ in which the arguments of the ϕ_{p_i} belong to the minimizing surface makes the determinant (1.5) always different from zero because of hypothesis (c). Since however ψ^* does not have the stated properties of differentiability, we may secure the desired function ψ by replacing the ϕ_{p_i} in ψ^* by polynomials $P_i(x,y)$ which approximate in the region B uniformly to the ϕ_{p_i} so closely that the determinant (1.5), calculated for the function $\psi = P_1 p_2 - P_2 p_1$, remains different from zero for (x,y) in B[1].

Admissible surfaces may be defined as surfaces given in the form (1.3) for which the functions $z_i(x,y)$ are of class C' in B and all of whose elements (x,y,z_i,p_i,q_i) lie interior to R and satisfy the equation (1.4). A sharper formulation, then, of the minimum properties of the surface under consideration is that this surface (1.1) gives to the integral (1.2) a value smaller than that furnished by any other admissible surface which passes through the same closed curve over the boundary of the region B.

We shall adopt the conventions of the tensor analysis, in which it is understood that an index s has always the same range $(s = 1,2,\dots,n)$; when the index appears twice in the same term, that term is to be regarded as a sum for values of s extending from 1 to n. The indices i,j,k in the following are always to have the ranges $(i,j,k = 1,2)$ while the index m

[1] E. W. Hobson, The Theory of Functions of a Real Variable, II(2nd ed., 1926), page 232.

has the range $(m = 1,2,3)$.

2. <u>The equation of variation</u>. Suppose $z_i = z_i(x,y,\epsilon)$ is a one-parameter family of admissible surfaces which for $\epsilon = 0$ contains the minimizing surface (1.1); we shall suppose that the functions $z_i(x,y,\epsilon)$ are of class C' for values (x,y,ϵ) having (x,y) in B and $|\epsilon|$ sufficiently small. When the functions $z_i(x,y,\epsilon)$ are substituted in equation (1.4) and the resulting identity is differentiated with respect to ϵ it is seen that the derivatives $\zeta_i(x,y) = z_{i\epsilon}(x,y,0)$ satisfy the equation

$$\phi_{z_i}\zeta_i + \phi_{p_i}\zeta_{ix} = 0, \tag{2.1}$$

in which the derivatives of ϕ have arguments belonging to the surface (1.1). Equation (2.1) is the so-called <u>equation of variation</u> along the minimizing surface.

A set of <u>admissible variations</u> along the minimizing surface consists of two functions $\zeta_i(x,y)$ which with $\zeta_{ix}(x,y)$ are of class C' in B and satisfy the equation (2.1). Each pair of admissible variations $\zeta_i(x,y)$ along the surface (1.1) defines a function $\rho(x,y)$ of class C' in B:

$$\rho(x,y) = \psi_{z_i}\zeta_i + \psi_{p_i}\zeta_{ix}. \tag{2.2}$$

The arguments in the derivatives of ψ in equation (2.2) are again those belonging to the minimizing surface. Conversely, each function $\rho(x,y)$ of class C' in B defines uniquely by means of equations (2.1) and (2.2) a pair of admissible variations $\zeta_i(x,y)$ with the initial values $\zeta_i[x_1(y),y] = 0$ on the interval y_1y_2.

Let now $\zeta_{mi}(x,y)$ be three pairs of admissible variations

along the surface (1.1), with $\zeta_{mi}[x_1(y), y] = 0$ on the interval $y_1 y_2$. Consider the equations

$$(2.3) \qquad \begin{aligned} \phi(x, y, z_i, p_i) &= 0, \\ \psi(x, y, z_i, p_i) &= r(x, y) + \epsilon_m \rho_m(x, y), \end{aligned}$$

where the functions ρ_m are defined from the variations ζ_{mi} by equation (2.2), and $r(x,y) = \psi[x, y, Z_i(x,y), Z_{ix}(x,y)]$. Equations (2.3) are ordinary differential equations of the first order in the functions z_i and their x-derivatives p_i and contain the parameters y and ϵ_m. The functional determinant of the left sides of (2.3) as to the p_i is (1.5), which does not vanish at any point of the surface (1.1), and for $\epsilon_m = 0$ $(m = 1,2,3)$ the functions $z_i = Z_i(x,y)$ are a solution of (2.3). According to the existence theorems for differential equations, then, the equations (2.3) have a unique solution $z_i = z_i(x, y, \epsilon_m)$ having the initial values

$$(2.4) \qquad z_i[x_1(y), y, \epsilon_m] = Z_i[x_1(y), y],$$

which reduces to the solution $Z_i(x,y)$ when the ϵ_m all vanish. The functions $z_i(x, y, \epsilon_m)$ and their derivatives $z_{ix}(x, y, \epsilon_m)$ are of class C' for values of their arguments in a sufficiently small neighborhood of the sets (x, y, ϵ_m) satisfying the conditions: (x,y) in B, $\epsilon_m = 0$. The functions $z_{i\epsilon_m}(x, y, 0)$ satisfy equations (2.1) and (2.2) written with the functions ρ_m, as one readily sees upon substituting the functions $z_i(x, y, \epsilon_m)$ in equations (2.3), differentiating the resulting identities with respect to ϵ_m, and setting the ϵ_m all equal to zero; from (2.4)

one sees that the functions $z_{1\epsilon_m}(x,y,0)$ satisfy also the initial conditions $z_{1\epsilon_m}[x_1(y),y,0] = 0$ $(y_1 \leq y \leq y_2)$; therefore $z_{1\epsilon_m}(x,y,0) = \zeta_{mi}(x,y)$, since the latter are the unique solutions of (2.1) and (2.2) vanishing at $x = x_1(y)$.

THEOREM. _If $\zeta_{mi}(x,y)$ are three pairs of admissible variations along $z_1 = Z_1(x,y)$, vanishing at $x = x_1(y)$, there exists a three-parameter family of admissible surfaces $z_1 = z_1(x,y,\epsilon_m)$ which contains the minimizing surface for $\epsilon_m = 0$ and has the functions $\zeta_{mi}(x,y)$ as its variations with respect to ϵ_m. The functions $z_1(x,y,\epsilon_m)$ and their derivatives z_{1x} are of class C' for values of (x,y,ϵ_m) in the neighborhood of the sets satisfying the conditions: (x,y) in B, $\epsilon_m = 0$, and satisfy the conditions : $z_1[x_1(y),y,\epsilon_m] = Z_1[x_1(y),y]$._

3. _An auxiliary minimum problem._ The theorem of the preceding section is still true if for the parameters ϵ_m we introduce functions $\epsilon_m(y)$ of class C' on the interval y_1y_2, on account of the fact that the equation (1.4) contains the derivatives p_1 and not the derivatives q_1. The family of admissible surfaces constructed with the aid of these functions $\epsilon_m(y)$, when substituted in the integral I, gives to it the form

$$I = \int_{y_1}^{y_2} g(y,\epsilon_m,\epsilon_m')\,dy, \tag{3.1}$$

where we have put

$$g(y,\epsilon_m,\epsilon_m') = \int_{x_1(y)}^{x_2(y)} f(x,y,z_1,z_{1x},z_{1y} + z_{1\epsilon_m}\epsilon_m')\,dx. \tag{3.2}$$

Then, by virtue of (2.4) and the discussion on page 8, the family $z_1(x,y,\epsilon_m)$ satisfies the conditions

$$z_1[x_1(y),y,\epsilon_m(y)] = Z_1[x_1(y),y] \quad (y_1 \leqslant y \leqslant y_2),$$
$$z_1[x,y_1,\epsilon_m(y_1)] = Z_1(x,y_1) \quad [x_1(y_1) \leqslant x \leqslant x_2(y_1)],$$
$$z_1[x,y_2,\epsilon_m(y_2)] = Z_1(x,y_2) \quad [x_1(y_2) \leqslant x \leqslant x_2(y_2)].$$

If the functions $\epsilon_m(y)$ are so chosen that in addition

$$z_1[x_2(y),y,\epsilon_m(y)] = Z_1[x_2(y),y] \quad (y_1 \leqslant y \leqslant y_2)$$

then each surface of the family will coincide with the minimizing surface (1.1) on the boundary of B. Since for $\epsilon_m(y) \equiv 0$ the family reduces to the minimizing surface, the following statement is justified.

<u>The functions $\epsilon_m(y) \equiv 0$ furnish a minimum for the integral</u>

$$I = \int_{y_1}^{y_2} g(y,\epsilon_m,\epsilon_m')\,dy$$

<u>in the class of sets of functions $\epsilon_m(y)$ of class C' on the interval y_1y_2 having all their elements $(y,\epsilon_m,\epsilon_m')$ in a small enough neighborhood of the sets $(y,0,0)$ with y on the interval y_1y_2 and satisfying the conditions</u>

(3.3) $$\epsilon_m(y_1) = \epsilon_m(y_2) = 0,$$

(3.4) $$z_1[x_2(y),y,\epsilon_m(y)] = Z_1[x_2(y),y] \quad (y_1 \leqslant y \leqslant y_2).$$

The auxiliary minimum problem so derived is a problem of Lagrange for simple integrals in (y,ϵ_m)-space, with equations of condition (3.4) which do not contain the derivatives

ϵ_m'. If the minimizing set $\epsilon_m(y) \equiv 0$ for this problem is non-singular - i.e., if the matrix

$$(3.5) \qquad \left\| z_{i\epsilon_m}[x_2(y), y, 0] \right\| = \left\| \zeta_{mi}[x_2(y), y] \right\|$$

has rank two at every point of the interval y_1y_2 - then it is well-known[1] that there exist multipliers $\mu_i(y)$ of class C' on the interval y_1y_2, such that the equations

$$(3.6) \qquad (d/dy)G_{\epsilon_m'} - G_{\epsilon_m} = 0,$$

formed for the function

$$(3.7) \quad G = g(y, \epsilon_m, \epsilon_m') + \mu_i(y)\left\{z_i[x_2(y), y, \epsilon_m] - Z_i[x_2(y), y]\right\},$$

hold along $\epsilon_m(y) \equiv 0$.

4. The multiplier rule for normal minimizing surfaces. The minimizing surface $z_i = Z_i(x, y)$ will be said to be normal if along it there exist two pairs of admissible variations $\zeta_{ki}(x, y)$ vanishing at $x = x_1(y)$, for which the determinant

$$(4.1) \qquad \begin{vmatrix} \zeta_{11}[x_2(y), y] & \zeta_{12}[x_2(y), y] \\ \zeta_{21}[x_2(y), y] & \zeta_{22}[x_2(y), y] \end{vmatrix}$$

never vanishes on the interval y_1y_2.

Suppose then that the minimizing surface is normal. If the two pairs of variations ζ_{ki} which satisfy the above condition and any third pair $\zeta_{3i}(x, y)$ which also vanish at $x = x_1(y)$

[1] See G. A. Bliss, The Problem of Lagrange in the Calculus of Variations, American Journal of Mathematics, XLII(1930), pages 703-5.

are employed in constructing the family $z_i(x,y,\epsilon_m)$, then the matrix (3.5) has surely rank two at every point of y_1y_2 and there exist multipliers $\mu_i(y)$ of class C' on the interval such that along $\epsilon_m(y) \equiv 0$ the equations (3.6) hold. These have the form

$$(d/dy)\int_{x_1(y)}^{x_2(y)} f_{q_i}\zeta_{mi}dx$$

(4.2)
$$= \mu_i\zeta_{mi}[x_2(y),y] + \int_{x_1(y)}^{x_2(y)} (f_{z_i}\zeta_{mi} + f_{p_i}\zeta_{mix} + f_{q_i}\zeta_{miy})dx,$$

the arguments in the derivatives of f belonging to the surface (1.1). The multipliers $\mu_i(y)$, which apparently depend upon the choice of the variations $\zeta_{si}(x,y)$, are in fact completely determined by the first two of equations (4.2), since the determinant (4.1) is not zero. Hence the equation (4.2) must hold when the $\zeta_{mi}(x,y)$ are replaced by an arbitrary pair of admissible variations $\zeta_i(x,y)$ which vanish at $x = x_1(y)$.

Let $\lambda_i(x,y)$ be two arbitrary functions of class C' in B. Then if $\rho(x,y)$ is determined from the variations $\zeta_i(x,y)$ by equation (2.2), we have

$$(d/dy)\int_{x_1(y)}^{x_2(y)} F_{q_i}\zeta_i dx - \int_{x_1(y)}^{x_2(y)} (F_{z_i}\zeta_i + F_{p_i}\zeta_{ix} + F_{q_i}\zeta_{iy})dx$$
$$= \mu_i\zeta_i[x_2(y),y] - \int_{x_1(y)}^{x_2(y)} \lambda_2\rho dx,$$

where we have put $F = f + \lambda_1\phi + \lambda_2\psi$. When the second term on the left of the last equation has been subjected to an integration by parts, it follows from the rule for differentiation under the integral sign and from $\zeta_1[x_1(y),y] \equiv 0$ that

$$\left[F_{q_1}(dx/dy) - F_{p_1} - \mu_1\right]\zeta_1\Big|^{x_2(y)}$$

$$- \int_{x_1(y)}^{x_2(y)} \left[F_{z_1} - (\partial/\partial x)F_{p_1} - (\partial/\partial y)F_{q_1}\right]\zeta_1 dx = - \int_{x_1(y)}^{x_2(y)} \lambda_2\rho\, dx,$$

the notation which follows the first term indicating that in it one must put $x = x_2(y)$.

If the functions $\lambda_1(x,y)$ are chosen as solutions of the ordinary linear differential equations

$$F_{z_1} - (\partial/\partial x)F_{p_1} - (\partial/\partial y)F_{q_1} = 0 \tag{4.3}$$

with the initial conditions

$$\left[F_{p_1} + \mu_1 - F_{q_1}(dx/dy)\right]^{x_2(y)} = 0, \tag{4.4}$$

which are linear in the functions λ_1, it follows that

$$\iint_B \lambda_2\rho dxdy = 0 \tag{4.5}$$

for arbitrary functions $\rho(x,y)$ of class C' in B. Hence $\lambda_2(x,y)$ is identically zero in B. For if λ_2 is different from zero at some point (x_0,y_0) of B, because of its continuity it will remain different from zero in a δ-neighborhood of that point. It would then be possible to find a function $\rho(x,y)$ which within

$(x_o, y_o)_\delta$ has the same sign as $\lambda_2(x,y)$ and vanishes outside this neighborhood. For this function $\rho(x,y)$ however the integral (4.5) is not zero.

THEOREM. *If $z_1 = Z_1(x,y)$ is a normal surface of class C'' which furnishes a minimum for the integral*

$$I = \iint_B f(x,y,z_1,p_1,q_1)\,dxdy$$

in the class of surfaces $z_1 = z_1(x,y)$ of class C' having the same boundary curve as $z_1 = Z_1(x,y)$ over the boundary of B and satisfying the equation $\phi(x,y,z_1,p_1) = 0$, and if along the surface $z_1 = Z_1(x,y)$ the derivatives ϕ_{p_1}, ϕ_{p_2} never vanish simultaneously, then there exists a function $\lambda(x,y)$ of class C' in B such that the equations

$$F_{z_1} - (\partial/\partial x)F_{p_1} - (\partial/\partial y)F_{q_1} = 0, \tag{4.6}$$

formed for the function $F = f + \lambda\phi$, hold along $z_1 = Z_1(x,y)$.

5. *Uniqueness of the multiplier for normal minimizing surfaces.* The following theorem is readily established.

THEOREM. *If the minimizing surface $z_1 = Z_1(x,y)$ is normal, the equations (4.6), formed for the function $F = \lambda_o f + \lambda\phi$, can hold along it for no pair of multipliers λ_o, $\lambda(x,y)$ having $\lambda_o = 0$, except the trivial pair which has also $\lambda(x,y) \equiv 0$.*

Suppose that the equations

$$\lambda\phi_{z_1} - (\partial/\partial x)(\lambda\phi_{p_1}) = 0 \tag{5.1}$$

are satisfied along the normal minimizing surface $z_1 = Z_1(x,y)$.

If the equations (5.1) are multiplied by the functions $\zeta_{ki}(x,y)$ which satisfy the equation (2.1) and have their determinant (4.1) different from zero, it is seen that $(\partial/\partial x)(\lambda\phi_{p_i}\zeta_{ki}) \equiv 0$ and that therefore for $x = x_2(y)$ the functions $\lambda\phi_{p_i}\zeta_{ki}$ vanish on the interval y_1y_2. It follows from the fact that (4.1) is not zero that $\lambda\phi_{p_i} \equiv 0$ for $x = x_2(y)$ and therefore $\lambda[x_2(y), y] \equiv 0$. But $\lambda(x,y)$ is a solution of

$$\lambda\left\{\left[\phi_{z_1} - (\partial/\partial x)\phi_{p_1}\right]\phi_{p_1} + \left[\phi_{z_2} - (\partial/\partial x)\phi_{p_2}\right]\phi_{p_2}\right\} - (\phi_{p_1}{}^2 + \phi_{p_2}{}^2)(\partial/\partial x)\lambda = 0$$

and hence $\lambda(x,y) \equiv 0$.

THEOREM. <u>The multiplier $\lambda(x,y)$ for the normal minimizing surface of the theorem on page 14 is unique.</u>

For if $\lambda_1(x,y)$, $\lambda_2(x,y)$ are two multipliers associated with the surface $z_i = Z_i(x,y)$, the equations (4.6), formed for the function $F = (\lambda_1 - \lambda_2)\phi$, must hold along $z_i = Z_i(x,y)$; but by the preceding theorem this is impossible unless $\lambda_1 \equiv \lambda_2$.

CHAPTER II

A MORE GENERAL PROBLEM

6. Formulation. The methods of the preceding sections are applicable also to problems of a more general type. Again let the surface $z_1 = Z_1(x,y)$ be of class C'' in B and suppose it gives a minimum to the integral I in the class of admissible surfaces which pass through the same closed curve as the surface (1.1) above the boundary of B - where now a surface is admissible if

(a) it is defined by functions $z_i = z_i(x,y)$ of class C' in B;

(b) all its elements (x,y,z_i,z_{ix},z_{iy}) are interior to the region R in which f and ϕ are of class C'' and to which the elements of the minimizing surface are all interior;

(c) all its elements satisfy the partial differential equation

$$\phi(x,y,z_1,z_2,p_1,p_2,q_2) = 0. \tag{6.1}$$

In equation (6.1) appear all the first derivatives of the functions $z_i(x,y)$ except z_{1y}. We replace the assumption (c) of the first section by a less symmetrical one; namely, that ϕ_{p_1} and ϕ_{q_2} are both different from zero on the surface $z_1 = Z_1(x,y)$. The assumption $\phi_{q_2} \neq 0$ implies that the case here considered is not that treated in the preceding chapter.

7. <u>The equations of variation</u>. If $z_2(x,y)$ is replaced in equation (6.1) by $Z_2(x,y) + \epsilon(y)\zeta_2(x,y)$, $\epsilon(y)$ and $\zeta_2(x,y)$ being arbitrary functions of their arguments, then z_1 can be determined from the equations

$$\begin{aligned} &\phi(x,y,z_1,\ Z_2 + \epsilon\zeta_2\ ,z_{1x},\ Z_{2x} + \epsilon\zeta_{2x},\ Z_{2y} + \epsilon\zeta_{2y} + \epsilon'\zeta_2) = 0, \\ (7.1)& \\ &z_1[x_1(y),y,\epsilon,\epsilon'] = Z_1[x_1(y),y], \end{aligned}$$

where ϵ' is the derivative of $\epsilon(y)$. If in these equations ϵ, ϵ' are regarded for the present as constant parameters, the solution will be a function $z_1(x,y,\epsilon,\epsilon')$ whose variations with respect to ϵ and ϵ', $\zeta_1 = z_{1\epsilon}(x,y,0,0)$ and $\zeta_1^* = z_{1\epsilon'}(x,y,0,0)$ satisfy the equations

$$(7.2) \qquad \phi_{z_1}\zeta_1 + \phi_{z_2}\zeta_2 + \phi_{p_1}\zeta_{1x} + \phi_{p_2}\zeta_{2x} + \phi_{q_2}\zeta_{2y} = 0,$$

$$(7.3) \qquad \phi_{z_1}\zeta_1^* + \phi_{p_1}\zeta_{1x}^* + \phi_{q_2}\zeta_2 = 0,$$

where as before the arguments in the derivatives of ϕ in these equations belong to the surface (1.1). Every function $\zeta_2(x,y)$ of class C'' in B determines uniquely a pair of functions ζ_1, ζ_1^* which with ζ_{1x}, ζ_{1x}^* are of class C' in B and which are solutions of the equations (7.2) and (7.3) respectively with the initial conditions

$$(7.4) \qquad \zeta_1[x_1(y),y] = \zeta_1^*[x_1(y),y] = 0 \quad (y_1 \leq y \leq y_2).$$

For use in the sequel, it should be noted that if for some y_0 on the interval y_1y_2, $\zeta_2(x,y_0) = \zeta_{2y}(x,y_0) \equiv 0$, then the functions $\zeta_1(x,y_0)$, $\zeta_1^*(x,y_0)$, $\zeta_{1y}^*(x,y_0)$ also vanish identically.

For from $\zeta_2(x,y_0) \equiv 0$ it follows that $\zeta_{2x}(x,y_0) \equiv 0$, and then the equation (7.2) reduces to

$$\phi_{z_1}\zeta_1 + \phi_{p_1}\zeta_{1x}\Big|^{y_0} = 0,$$

of which the only solution with initial value zero is $\zeta_1(x,y_0)$ identically zero. Similarly from (7.3) one finds $\zeta_1^*(x,y_0) \equiv 0$. Differentiating (7.3) with respect to y and observing that $\zeta_{1x}^*(x,y_0) \equiv 0$ since $\zeta_1^*(x,y_0) \equiv 0$, one derives the equation

$$\phi_{z_1}\zeta_{1y}^* + \phi_{p_1}\zeta_{1yx}^*\Big|^{y_0} = 0.$$

From (7.4) one sees that the solution $\zeta_{1y}^*(x,y_0)$ of this equation must satisfy the initial condition

$$\zeta_{1y}^*\left[x_1(y_0),y_0\right] = (d/dy)\,\zeta_1^*\left[x_1(y_0),y_0\right] - \zeta_{1x}^*\left[x_1(y_0),y_0\right]\,x_1'(y_0) = 0$$

from which it follows that $\zeta_{1y}^*(x,y_0) \equiv 0$.

THEOREM. <u>If $\zeta_{m2}(x,y)$ are three arbitrary functions of class C'' in B for which</u>

$$(7.5)\qquad \zeta_{m2}\left[x_1(y),y\right] = \zeta_{m2}\left[x_2(y),y\right] = 0 \quad (y_1 \leq y \leq y_2),$$

<u>then there exists a pair of functions</u>

$$z_1 = z_1(x,y,\epsilon_m,\epsilon_m'),$$

$$z_2 = z_2(x,y,\epsilon_m) = Z_2(x,y) + \epsilon_m\zeta_{m2}(x,y),$$

<u>containing arbitrary constants ϵ_m, ϵ_m', and satisfying the equations (7.1) with the initial conditions</u>

$$z_1\left[x_1(y),y,\epsilon_m,\epsilon_m'\right] = Z_1\left[x_1(y),y\right].$$

The family contains the minimizing surface $z_1 = Z_1(x,y)$ for the values $(x,y,\epsilon_m,\epsilon_m')$ satisfying the conditions

$$(x,y) \text{ in } B, \quad \epsilon_m = \epsilon_m' = 0.$$

In a neighborhood of these values $(x,y,\epsilon_m,\epsilon_m')$ the function $z_1(x,y,\epsilon_m,\epsilon_m')$ and its derivative z_{1x} are of class C'. Furthermore the variations $\zeta_{mi} = z_{i\epsilon_m}$, $\zeta_{mi}^* = z_{i\epsilon_m'}$ are those which are determined uniquely by the equations (7.2), (7.3) and the initial conditions (7.4) when the functions $\zeta_{m2}(x,y)$ are given. If ϵ_m, ϵ_m' are replaced by functions $\epsilon_m(y)$ of class C'' on y_1y_2 and their derivatives $\epsilon_m' = d\epsilon_m/dy$, then the surface

$$
\begin{aligned}
z_1 &= z_1\left[x,y,\epsilon_m(y),\epsilon_m'(y)\right], \\
z_2 &= z_2\left[x,y,\epsilon_m(y)\right]
\end{aligned}
\tag{7.6}
$$

still satisfies the equation $\phi = 0$ and is admissible.

The statements made in this theorem follow from a consideration of the differential equation (7.11) in which the ϵ_m, ϵ_m' and y are parameters. Since $\phi_{p_1} \neq 0$ along the minimizing surface, the existence theorems for differential equations assure the existence of a function z_1 of the kind described; the equations $\zeta_{mi} = z_{i\epsilon_m}$, $\zeta_{mi} = z_{i\epsilon_m'}$ follow from substitution of the functions $z_1(x,y,\epsilon_m,\epsilon_m')$, $z_2(x,y,\epsilon_m)$ in the equation (7.1) and differentiation with respect to ϵ_m, ϵ_m', since ζ_{mi}, ζ_{mi}^* are the unique solutions of the equations (7.2), (7.3) respectively which for each m have the initial values (7.4). The truth of the last statement of the theorem is evident if one examines the form of equation (7.1).

8. A special case of the multiplier rule for simple integrals. If the functions $\epsilon_m(y)$ of the theorem of section 7 are chosen to satisfy the equations

$$(8.1) \qquad \epsilon_m(y_1) = \epsilon_m(y_2) = \epsilon_m'(y_1) = \epsilon_m'(y_2) = 0,$$

$$(8.2) \qquad z_1\left[x_2(y), y, \epsilon_m, \epsilon_m'\right] - Z_1\left[x_2(y), y\right] = 0,$$

the surfaces of the family (7.6) will coincide with the minimizing surface above the entire boundary of B. Because the family reduces to the surface (1.1) for $\epsilon_m(y) \equiv 0$, the following statement is justified:

If the surface $z_1 = Z_1(x,y)$ minimizes the integral I, then the functions $\epsilon_m(y) \equiv 0$ must give to the integral

$$(8.3) \qquad J = \int_{y_1}^{y_2} g(y, \epsilon_m, \epsilon_m')\, dy,$$

where

$$(8.4) \qquad g = \int_{x_1(y)}^{x_2(y)} f\left[x, y, z_1(x,y,\epsilon_m,\epsilon_m'), z_2(x,y,\epsilon_m), z_{1x}, z_{2x}, z_{1y}, z_{2y}\right] dx,$$

a minimum value for the problem of minimizing J in the class of sets of functions $\epsilon_m(y)$ of class C'' on y_1y_2 for which the sets $(y, \epsilon_m, \epsilon_m')$ are near to $(y, 0, 0)$ and which satisfy the end-conditions (8.1) and the equations (8.2).

The auxiliary minimum problem so derived is again a problem of Lagrange for simple integrals. It differs, however, from the type of problem usually considered in that not only the functions $\epsilon_m(y)$ but their derivatives $\epsilon_m'(y)$ as well are subjected to conditions at the ends of the interval $y_1 \leq y \leq y_2$.

However, the reduction of this problem to the classical type is easily carried out and will occupy us in the present section.

Replace the ϵ_m' by new variables η_m. Then the auxiliary minimum problem derived above is equivalent to that of minimizing the integral

$$J = \int_{y_1}^{y_2} g(y, \epsilon_m, \eta_m)\,dy$$

in the class of functions $\epsilon_m(y)$, $\eta_m(y)$ of class C' on y_1y_2 and satisfying the conditions $\epsilon_m(y_1) = \epsilon_m(y_2) = \eta_m(y_1) = \eta_m(y_2) = 0$ and the equations

$$(8.5) \qquad \begin{gathered} z_1\left[x_2(y), y, \epsilon_m, \eta_m\right] - z_1\left[x_2(y), y\right] = 0, \\ \epsilon_m' - \eta_m = 0. \end{gathered}$$

This is an instance of a problem with "mixed" equations of condition, since the first of equations (8.5) does not contain the derivatives ϵ_m' or η_m'. The set $\epsilon_m(y) = \eta_m(y) \equiv 0$ is a solution of the problem, and it is known[1] that if one of the fourth order determinants of the matrix

$$(8.6) \qquad \left\| \begin{matrix} \zeta_{11} & \zeta_{21} & \zeta_{31} & \zeta_{11}^* & \zeta_{21}^* & \zeta_{31}^* \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{matrix} \right\|_{x_2(y)}$$

of the derivatives of the first equation (8.5) with respect to ϵ_m, η_m and of the last three with respect to ϵ_m', η_m' is never

---

[1] Bolza, Vorlesungen über Variationsrechnung, page 580.

zero, then there exist multipliers λ_o, $\mu(y)$, $\nu_m(y)$ of class C' on y_1y_2 and not all zero, such that the equations

$$(8.7)\qquad \begin{aligned} &(d/dy)G_{\epsilon'_m} - G_{\epsilon_m} = 0, \\ &(d/dy)G_{\eta'_m} - G_{\eta_m} = 0 \end{aligned}$$

hold along $\epsilon_m(y) = \eta_m(y) \equiv 0$, where

$$G = \lambda_o g(y,\epsilon_m,\eta_m) + \mu\Big\{z_1\big[x_2(y),y,\epsilon_m,\eta_m\big] - z_1\big[x_2(y),y\big]\Big\} + \nu_m(\epsilon'_m - \eta_m).$$

For $\epsilon_m(y) = \eta_m(y) \equiv 0$, the equations (8.7) are equivalent to

$$(8.8)\qquad \begin{aligned} &(d/dy)\Big\{\lambda_o g_{\epsilon'_m}(y,0,0) + \mu\zeta^*_{m1}\big[x_2(y),y\big]\Big\} \\ &\qquad\qquad = \lambda_o g_{\epsilon_m}(y,0,0) + \mu\zeta_{m1}\big[x_2(y),y\big], \end{aligned}$$

which are precisely the equations one would arrive at in calculating

$$(d/dy)G_{\epsilon'_m}(y,0,0) - G_{\epsilon_m}(y,0,0) = 0$$

for the function

$$G(y,\epsilon_m,\epsilon'_m) = \lambda_o g(y,\epsilon_m,\epsilon'_m) + \mu\Big\{z_1\big[x_2(y),y,\epsilon_m,\epsilon'_m\big] - z_1\big[x_2(y),y\big]\Big\}.$$

9. <u>The multiplier rule for normal minimizing surfaces.</u> In the problem under consideration we shall call an admissible surface <u>normal</u> in case there exist two functions $\zeta_{12}(x,y)$ of class C'' vanishing for $x = x_1(y)$ and $x = x_2(y)$ and such that for the functions $\zeta_{11}(x,y)$, $\zeta^*_{11}(x,y)$ associated with them one of the values $\zeta^*_{11}[x_2(y),y]$ is always different from zero on the interval y_1y_2, while the determinant

$$(9.1)\qquad \left| \zeta_{1i}^{*}[x_2(y),y] \qquad (d/dy)\,\zeta_{1i}^{*}[x_2(y),y] - \zeta_{1i}[x_2(y),y] \right|$$

never vanishes on the interval y_1y_2.

Suppose then that the minimizing surface $z_1 = Z_1(x,y)$ is normal. Then according to Bolza there exist multipliers λ_0, $\mu(y)$ for which equations (8.8) hold. These equations, when written in the form

$$(9.3)\qquad \begin{aligned} &\mu'\zeta_{mi}^{*}[x_2(y),y] + \mu\left\{(d/dy)\,\zeta_{mi}^{*}[x_2(y),y] - \zeta_{mi}[x_2(y),y]\right\} \\ &= \lambda_0\left\{\int_{x_1(y)}^{x_2(y)}(f_{z_1}\zeta_{mi} + f_{p_1}\zeta_{mix} + f_{q_1}\zeta_{miy})\,dx \right. \\ &\qquad \left. - (d/dy)\int_{x_1(y)}^{x_2(y)}(f_{z_1}\zeta_{mi}^{*} + f_{p_1}\zeta_{mix}^{*} + f_{q_1}\zeta_{miy}^{*} + f_{q_2}\zeta_{m2})\,dx\right\}, \end{aligned}$$

may be regarded as three linear equations having a non-trivial solution λ_0, $\mu(y)$, $\mu'(y)$, with the constant λ_0 different from zero. If λ_0 were zero, the fact that the determinant (9.1) is not zero on y_1y_2 would imply that $\mu(y) \equiv 0$ and then equations (8.7) would show that also the $\nu_m(y) \equiv 0$, in contradiction with the multiplier rule for simple integrals referred to in the footnote on page 21. The solution of the auxiliary minimum problem is therefore normal[1], and the multipliers λ_0, $\mu(y)$ may be taken in the form $1, \mu(y)$, in which form they are unique.

---

[1] See Bliss, loc. cit., page 688.

Since the first two equations (9.3) suffice to determine the functions μ, μ' completely in view of the fact that the determinant (9.1) is not zero, it follows that equation (9.3) is still true when the functions ζ_{m2}, ζ_{m1}, ζ_{m1}^{*} are replaced by an arbitrary function $\zeta_2(x,y)$ of class C'', vanishing at $x = x_1(y)$ and $x = x_2(y)$, and its associated functions $\zeta_1(x,y)$, $\zeta_1^{*}(x,y)$.

Let $\lambda(x,y)$ be an arbitrary function of class C' in B. As a consequence of equations (7.2) and (7.3) one has

$$\mu'\zeta_1^{*}[x_2(y),y] + \mu\left\{(d/dy)\zeta_1^{*}[x_2(y),y] - \zeta_1[x_2(y),y]\right\}$$
$$= \int_{x_1(y)}^{x_2(y)} (F_{z_1}\zeta_1 + F_{p_1}\zeta_{1x} + F_{q_1}\zeta_{1y})\,dx$$
$$- (d/dy)\int_{x_1(y)}^{x_2(y)} (F_{z_1}\zeta_1^{*} + F_{p_1}\zeta_{1x}^{*} + F_{q_1}\zeta_{1y}^{*} + F_{q_2}\zeta_2)\,dx,$$

where $F = f + \lambda\phi$. When the terms on the right are integrated by parts and the conditions $\zeta_2[x_1(y),y] = \zeta_2[x_2(y),y] \equiv 0$ and (7.4) are recalled, one has

$$\mu'\zeta_1^{*}[x_2(y),y] + \mu\left\{(d/dy)\zeta_1^{*}[x_2(y),y] - \zeta_1[x_2(y),y]\right\}$$
(9.4)
$$= F_{p_1}\zeta_1\Big|^{x_2(y)} + \int_{x_1(y)}^{x_2(y)}\left[L_1\zeta_1 + (\partial/\partial y)(F_{q_1}\zeta_1)\right]dx$$
$$- (d/dy)\left\{F_{p_1}\zeta_1^{*}\Big|^{x_2(y)} + \int_{x_1(y)}^{x_2(y)}\left[L_1\zeta_1^{*} + (\partial/\partial y)(F_{q_1}\zeta_1^{*})\right]dx\right\},$$

where we have put $L_1 = F_{z_1} - (\partial/\partial x)F_{p_1} - (\partial/\partial y)F_{q_1}$. Let $\lambda(x,y)$ be chosen as a solution of the ordinary differential equation

$$F_{z_1} - (\partial/\partial x)F_{p_1} - (\partial/\partial y)F_{q_1} = 0 \tag{9.5}$$

with the initial condition

$$\Big[\mu + F_{p_1} - F_{q_1}(dx/dy)\Big]^{x_2(y)} = 0; \tag{9.6}$$

both these equations are linear, (9.6) in λ, and (9.5) in λ and $\partial\lambda/\partial x$. Then the equation (9.4), which when $L_1 = 0$ can be written in the form

$$K\zeta_1\Big|^{x_2(y)} - (d/dy)\Big\{K\big[x_2(y),y\big]\,\zeta_1^*\big[x_2(y),y\big]\Big\} + \int_{x_1(y)}^{x_2(y)} L_2\zeta_2\,dx$$

$$- \frac{d}{dy}\left\{\int_{x_1(y)}^{x_2(y)}\left[F_{q_1}(\zeta_{1y}^* - \zeta_1) + \zeta_1^*\frac{\partial}{\partial y}F_{q_1}\right]dx + \left(F_{q_1}\frac{dx}{dy}\zeta_1^*\right)\Big|^{x_2(y)}\right\} = 0,$$

where we have put $K = \mu + F_{p_1} - F_{q_1}(dx/dy)$, becomes

$$\int_{x_1(y)}^{x_2(y)} L_2\zeta_2\,dx - \frac{d}{dy}\left\{\int_{x_1(y)}^{x_2(y)}\left[F_{q_1}(\zeta_{1y}^* - \zeta_1) + \zeta_1^*\frac{\partial}{\partial y}F_{q_1}\right]dx + \left(F_{q_1}\frac{dx}{dy}\zeta_1^*\right)\Big|^{x_2(y)}\right\} = 0, \tag{9.7}$$

which must hold for every function $\zeta_2(x,y)$ of class C'' in B and vanishing at $x = x_1(y)$ and $x = x_2(y)$. In particular, then, equation (9.7) must hold for such functions $\zeta_2(x,y)$ which with their derivatives ζ_{2y} vanish for $y = y_1$ and $y = y_2$ identically in x. When the left member of equation (9.7) is integrated

with respect to y from y_1 to y_2, the second term is seen to vanish, in accordance with the discussion on pages 17 and 18. Therefore the integral

$$\iint_B L_2 \zeta_2 dxdy = 0$$

for arbitrary functions $\zeta_2(x,y)$ of class C'' which vanish on the boundary of B and whose derivatives ζ_{2y} vanish for y_1 and y_2. It follows from an argument similar to that employed on page 13 that $L_2 \equiv 0$.

THEOREM. <u>Let $z_1 = Z_1(x,y)$ be a normal surface of class C'' along which neither ϕ_{p_1} nor ϕ_{q_2} vanishes; if the surface furnishes a minimum for the integral I in the class of surfaces $z_1 = z_1(x,y)$ of class C' which have all their elements interior to the region R, have the same boundary curve as $z_1 = Z_1(x,y)$ over the boundary of B, and satisfy the equation</u>

$$\phi(x,y,z_1,z_2,p_1,p_2,q_2) = 0,$$

<u>then there exists a multiplier $\lambda(x,y)$ of class C' in B such that the equations</u>

$$F_{z_1} - (\partial/\partial x)F_{p_1} - (\partial/\partial y)F_{q_1} = 0, \tag{9.8}$$

<u>formed for the function $F = f + \lambda\phi$, hold along $z_1 = Z_1(x,y)$.</u>

10. <u>Uniqueness of the multiplier for normal minimizing surfaces.</u> The following theorem is readily established:

THEOREM. <u>Along a normal minimizing surface th equations</u>

(9.8) <u>can hold for no function $F = \lambda\phi$ unless $\lambda(x,y) \equiv 0$.</u>

Suppose that along the minimizing surface the equations

$$
(10.1)\quad
\begin{aligned}
&\lambda\phi_{z_1} - (\partial/\partial x)(\lambda\phi_{p_1}) = 0,\\
&\lambda\phi_{z_2} - (\partial/\partial x)(\lambda\phi_{p_2}) - (\partial/\partial y)(\lambda\phi_{q_2}) = 0
\end{aligned}
$$

are satisfied. Since the minimizing surface is normal, there exist functions ζ_{k2} for whose associated functions ζ_{k1}, ζ_{k1}^* the determinant (9.1) does not vanish. From equations (10.1) and the equations of variation it follows that

$$
\begin{aligned}
&(\partial/\partial x)(\lambda\phi_{p_1}\zeta_{k1}) + (\partial/\partial y)(\lambda\phi_{q_2}\zeta_{k2}) = 0,\\
&-\lambda\phi_{q_2}\zeta_{k2} = (\partial/\partial x)(\lambda\phi_{p_1}\zeta_{k1}^*),
\end{aligned}
$$

so that

$$(\partial/\partial x)\Big[\lambda\phi_{p_1}\zeta_{k1} - (\partial/\partial y)(\lambda\phi_{p_1}\zeta_{k1}^*)\Big] = 0.$$

Therefore

$$(10.2)\qquad \lambda\phi_{p_1}(\zeta_{k1y}^* - \zeta_{k1}) + \zeta_{k1}^*(\partial/\partial y)(\lambda\phi_{p_1})\Big|^{x_2(y)} = 0.$$

If now $\lambda(x,y) \not\equiv 0$ there exists a value y_0 on y_1y_2 for which $\lambda\left[x_2(y_0), y_0\right] \neq 0$, and hence from (10.2) the determinant

$$(10.3)\qquad \begin{vmatrix} \zeta_{k1y}^* - \zeta_{k1} & \zeta_{k1}^* \end{vmatrix}$$

vanishes at $x = x_2(y_0)$, $y = y_0$. But

$$
\begin{aligned}
\zeta_{k1y}^*\left[x_2(y), y\right] &= (d/dy)\zeta_{k1}^*\left[x_2(y), y\right] - \zeta_{k1x}^*\left[x_2(y), y\right](dx_2/dy)\\
&= (d/dy)\zeta_{k1}^*\left[x_2(y), y\right] + M\left[x_2(y), y\right]\zeta_{k1}^*\left[x_2(y), y\right],
\end{aligned}
$$

where $M(x,y) = (\phi_{z_1}/\phi_{p_1})(dx/dy)$; this follows from equation (7.3) and the fact that $\zeta_{k2}\left[x_2(y), y\right] \equiv 0$. Therefore the

determinant (10.3) is equal to the determinant (9.1) at least for $x = x_2(y)$. But the latter determinant is never zero on the interval y_1y_2.

THEOREM. _The multiplier associated with the normal minimizing surface_ $z_1 = Z_1(x,y)$ _by the theorem of section 9 is unique_.

The proof of this corollary of the preceding theorem is exactly like that given on page 15.

THE CONDITION OF MAYER FOR DISCONTINUOUS SOLUTIONS OF THE LAGRANGE PROBLEM

BY

RALPH AUBRIE HEFNER

TABLE OF CONTENTS

INTRODUCTION

In his paper on Discontinuous Solutions in Space Problems of the Calculus of Variations Graves[1] gives a new formulation of the Jacobi-Caratheodory condition for minimizing curves with corners which uses no special assumptions, other than that the function Ω_0 does not vanish at the corners. His method is applicable to minimizing curves in any number of dimensions and greatly simplifies the discussion while leaving no exceptions uncared for.

The purpose of this paper is to extend his method to the Mayer condition for the so-called discontinuous solutions of the Lagrange problem in the calculus of variations. This has been done by means of the second variation and a new definition of conjugate point in terms of solutions of the generalized accessory equations for the discontinuous problem. No additional assumptions have been necessary, other than that the function Ω_0 does not vanish at the corners of the minimizing arc. It is also shown that no minimizing curve can have a cusp at which Ω_0 does not vanish. Finally it is shown that the general solution of the Euler-Lagrange equations can be used in determining the conjugate points.

---

1. American Journal of Mathematics, vol. 52(1930) p.1.

1. STATEMENT OF THE PROBLEM.

A few often repeated indices will have, throughout the paper the following ranges:

$$i, j = 1, \ldots, n \; ;$$

$$\alpha, \beta = 1, \ldots, m < n-1 \; ;$$

$$h = 1, \ldots, 2n-2.$$

Prime marks will indicate derivatives with respect to the independent variable t, while subscripts will indicate partial derivatives. It will be convenient to use (y, y') for $(y^1, \ldots, y^n, y^{1\prime}, \ldots, y^{n\prime})$ and corresponding abbreviations for other similar expressions where no loss in clearness is caused thereby. We shall make use of the well-known summation convention of tensor analysis, except that it will not be necessary to distinguish between contravariant and covariant indices. As an example of our notation we have to deal with the norm $y^{\lambda\prime}y^{\lambda\prime}(=\sum_{\lambda=0}^{n} y^{\lambda\prime 2})$ of a vector $y'(= y^{1\prime}, \ldots, y^{n\prime})$. We shall also find it convenient to say that a function $y(t)$ is of class D' on an interval (t_0, T) if it is continuous on the interval and if the interval can be divided into a finite number of sub-intervals on each of which the function has a continuous derivative which has a finite limit at each end of the sub-interval. A function $y(t)$ will be said to be of class D on an interval if the interval can be divided into a finite number of sub-intervals on each of which the function is continuous and approaches finite limits at the ends.

With the aid of these conventions the problem of Lagrange in the calculus of variations may be stated as that

of finding among the arcs, $y^i = y^i(t)$, $(t_0 \leq t \leq T)$, which join two given points y_0, Y, and satisfy a system of differential equations of the form

$$\varphi^\alpha(y, y') = 0, \tag{1:1}$$

one which minimizes the integral

$$I \equiv \int_{t_0}^{T} f(y, y')\,dt. \tag{1:2}$$

By defining an <u>admissible</u> <u>arc</u> as an arc, $y^i = y^i(t)$, $(t_0 \leq t \leq T)$, whose functions $y^i(t)$ are of class D' on (t_0, T) and satisfy the equations (1:1), our problem may be stated more explicitly as that of finding among the class of all admissible arcs which join two given points y_0, Y one which minimizes the integral (1:2). It is important to recall that if the $y^{i\prime}$ are the derivatives of the functions y^i defining an admissible arc then $y^{i\prime}y^{i\prime}$ never vanishes along the arc.

As we proceed with the analysis of this problem we shall use the following additional hypotheses and definitions:[2]

(a) The functions $y^i(t)$ defining a minimizing arc E_0 are admissible.

(b) In a neighborhood $\mathcal{R}$ of the values (y, y') defining the arc E_0 the functions f, φ^α have the following properties: They are continuous and have continuous partial derivatives up to and including those of the fourth order in all their arguments. Such functions are usually said to be of class C^{IV}. They are positively homogeneous of the first degree in y'.

---

2. Bliss, G.A., <u>The problem of Lagrange in the calculus of variations</u>, <u>American Journal of Mathematics, vol.52(1930) p.675</u>

That is,

$$f(y,ky') = kf(y,y')$$

$$\varphi^\alpha(y,ky') = k\varphi^\alpha(y,y'),\ k > 0.$$

(c) At every element (y,y') of E_0 the $m \times n$ dimensional matrix

(1:3) $$\left\| \varphi^\alpha_{y^{i\prime}} \right\|$$

has rank m.

If a one-parameter family of admissible arcs

$$y^i = y^i(t,b)$$

containing a particular admissible arc E_0 for $b = b_0$ is given, the functions $\eta^i(t) \equiv y^i_b(t,b_0)$ are called <u>variations of the family along</u> E_0.

<u>The equations of variation</u> for the arc E_0 are defined by the formula

(1:4) $$\Phi^\alpha(\eta,\eta') = \eta^i\varphi^\alpha_{y^i} + \eta^{i\prime}\varphi^\alpha_{y^{i\prime}} = 0,$$

where the arguments of the coefficients $\varphi^\alpha_{y^i}$, $\varphi^\alpha_{y^{i\prime}}$ are the functions $y^i(t)$, $y^{i\prime}(t)$ belonging to E_0.

A set of functions $\eta^i(t)$ of class D' satisfying equations (1:4) is called an <u>admissible set of variations</u>.

2. THE EULER-LAGRANGE MULTIPLIER RULE[3].

If we let

$$F(y,y',\lambda) \equiv \lambda^0 f + \lambda^\alpha(t)\,\varphi^\alpha,$$

then the multiplier rule may be stated as follows:

THEOREM 2:1. <u>For every minimizing arc</u> E_0 <u>there exists</u>

---

3. Eshleman, Joel David, <u>The Lagrange problem in parametric form in the calculus of variations</u>, typed dissertation, Chicago, 1922, pp. 7-9.

a set of constants C^i, λ^0 and a set of functions $\lambda^\alpha(t)$ of class D such that λ^0, λ^α are not all identically zero and the equations

$$F_{y^{i\prime}} = \int_{t_0}^{t} F_{y^i}\,dt + C^i$$

hold along E_0.

COROLLARY 2:1. THE EULER-LAGRANGE EQUATIONS. On every sub-arc between corners of a minimizing arc E_0 the differential equations

$$(2:1) \qquad \frac{d}{dt} F_{y^{i\prime}} = F_{y^i}, \qquad \varphi^\alpha = 0$$

must be satisfied.

COROLLARY 2:2. THE CORNER CONDITION. At every corner of a minimizing arc E_0 the condition

$$(2:2) \qquad F_{y^{i\prime}}(y, y'^{+}, \lambda^{+}) = F_{y^{i\prime}}(y, y'^{-}, \lambda^{-})$$

must be satisfied. Here and throughout the paper a small + or - sign written as a superscript signifies that the function thus designated is evaluated on the forward or backward side of the corner respectively.

COROLLARY 2:3. THE DIFFERENTIABILITY CONDITION. Near every point $t = t_0$ where the minimizing arc E_0 is of class C' and at which the determinant

$$(2:3) \qquad R = \begin{vmatrix} F_{y^{i\prime}y^{j\prime}} & \varphi^\alpha_{y^{i\prime}} & y^{i\prime} \\ \varphi^\beta_{y^{j\prime}} & 0 & 0 \\ y^{j\prime} & 0 & 0 \end{vmatrix}$$

is different from zero the functions $y^i(t)$ defining E_0 are of class C^{IV} and the multipliers $\lambda^\alpha(t)$ are of class C''', at least.

3. EQUIVALENCE OF THE HYPOTHESES $F_1 \neq 0$ AND $R \neq 0$.

The determinant that naturally arises when we seek to solve the Euler-Lagrange equations for their highest derivatives is the determinant

$$D = \begin{vmatrix} F_{y^{i\prime} y^{j\prime}} & \varphi^{\alpha}_{y^{i\prime}} \\ \varphi^{\alpha}_{y^{j\prime}} & 0 \end{vmatrix} .$$

However when the problem is in parametric form $D \equiv 0$, due to the homogeneity conditions assumed for the functions f, φ^{α}. The usual hypothesis in this case is to assume D of rank $n+m-1$, and we have the following theorems:

THEOREM 3:1. <u>If A^{rs} is the cofactor of the element in the r^{th} row and s^{th} column of the determinant D, $(r,s,=1,\ldots,n+m)$, then there exists a function F_1 such that $A^{rs} = p^r F_1 p^s$, where $p^i = y^{i\prime}$ and $p^{\gamma} = 0$, $(\gamma = n+1,\ldots,n+m)$.</u>

This theorem[4] is a consequence of the identities

$$(3:1) \qquad \begin{aligned} y^{i\prime} F_{y^{i\prime}} &= F, \\ y^{i\prime} \varphi^{\alpha}_{y^{i\prime}} &= \varphi^{\alpha}, \\ y^{i\prime} F_{y^{i\prime} y^{j\prime}} &= 0, \end{aligned}$$

which are easily derived from the homogeneity conditions assumed for the functions f, φ^{α}.

THEOREM 3:2. <u>If D is of rank $n+m-1$ along an admissible arc E_0, the determinant R in (2:3) is different from zero everywhere along the admissible arc E_0, and conversely.</u>

---

4. Bliss and Mason, <u>The properties of curves in space which minimize a definite integral.</u> Trans. Amer. Math. Soc. vol. 9(1908) p.241. Also Bliss, <u>A note on symmetric Matrices</u>, Annals of Mathematics, vol. 16(1914-15) p.43.

If we expand R by the last row and column and make use of theorem 3:1 it reduces to

$$R = -(y^{i\prime}y^{i\prime})^2 F_1$$

which is readily seen to be different from zero along E_0 under the hypothesis assumed.

With these theorems before us we see that if we take for our fundamental hypothesis that D is of rank n+m-1 it follows at once that $F_1 \neq 0$ and $R \neq 0$. Hence we can use any one of the three hypotheses that is most convenient in the discussions that follow.

4. THE EXTREMALS.

An admissible arc and set of multipliers

$$y^i = y^i(t), \quad \lambda^0, \qquad \lambda^\alpha = \lambda^\alpha(t), \qquad (t_0 \leq t \leq t_1),$$

is called an <u>extremal</u> if it has continuous derivatives $y^{i\prime}(t)$, $y^{i\prime\prime}(t)$, $\lambda^{\alpha\prime}(t)$ on the interval (t_0, t_1) and if it satisfies the Euler-Lagrange equations (2:1).

In order to obtain an imbedding theorem that will be useful in the sequel we reduce the system of differential equations (2:1) to a canonical form by means of the following device. We select the arc length as parameter and consider the system of equations

$$\begin{aligned} &\frac{d}{dt}F_{y^{i\prime}} - F_{y^i} + uy^{i\prime} = 0, \\ (4:1) \qquad &\frac{d}{dt}\varphi^\alpha = 0, \\ &\frac{d}{dt}(y^{i\prime}y^{i\prime}-1) = 0, \end{aligned}$$

and initial conditions

$$(4:2)\qquad \begin{aligned} &\varphi^{\alpha}(y(t_o),\, y'(t_o)) = 0, \\ &y^{i\prime}(t_o)y^{i\prime}(t_o) = 1. \end{aligned}$$

If we multiply the first n equations of (4:1) by $y^{i\prime}$ respectively and add we find, by making use of the identities (3:1) and initial conditions (4:2) that $u = 0$ is the only value of u satisfying equations (4:1) and (4:2). Hence the system of equations (4:1) together with the initial conditions (4:2) is entirely equivalent to the system of equations (2:1) plus the equation $y^{i\prime}y^{i\prime} = 1$. Also the system of equations (4:1) is linear in the variables $y^{j\prime\prime}$, $\lambda^{\beta\prime}$, u, and the determinant of their coefficients is R. Hence near an extremal E_o along which $R \neq 0$ these equations can be solved for $y^{j\prime\prime}$, $\lambda^{\beta\prime}$, u and they are readily seen to be equivalent to the system

$$(4:3)\qquad u = 0, \qquad \frac{dy^i}{dt} = y^{i\prime}, \qquad \frac{dy^{i\prime}}{dt} = G^i(y, y', \lambda), \qquad \frac{d\lambda^{\alpha}}{dt} = H^{\alpha}(y, y', \lambda)$$

in so called canonical form. To this latter system of equations we may apply the imbedding theorem of Bolza[5] and obtain the following theorem.

THEOREM 4:1. _If_ $y^i(t)$, $\lambda^{\alpha}(t)$, $(t_o \leq t \leq t_1)$ _is a solution of equations_ (4:3), _and if_ τ _be some special value of t in the interval_ (t_o, t_1) _making_

$$y^i(\tau) = \xi_o^i, \quad y^{i\prime}(\tau) = \xi_o^{i\prime}, \quad \lambda^{\alpha}(\tau) = \ell_o^{\alpha},$$

then for every system of values (ξ, ξ', ℓ) _sufficiently near to those of_ (y, y', λ) _there exists a solution_

---

5. _Vorlesungen uber Variationsrechnung_, p.179.

$$
(4:4) \qquad \begin{aligned} y^i &= \psi^i(t,\xi,\xi',\ell), \\ y^{i\prime} &= \psi^{i\prime}(t,\xi,\xi',\ell), \\ \lambda^\alpha &= L^\alpha(t,\xi,\xi',\ell), \end{aligned}
$$

of equations (4:3) throughout the interval (t_0, t_1). The solution is of class C'' and lies wholly within $\mathcal{R}$. For $\xi^i = \xi^i_0$, $\xi^{i\prime} = \xi^{i\prime}_0$, $\ell^\alpha = \ell^\alpha_0$, the solution (4:4) reduces identically to $y^i(t)$, $y^{i\prime}(t)$, $\lambda^\alpha(t)$, on the interval (t_0, t_1).

In order to go from the system of equations (4:3) back to the system (2:1) we find that m of our constants are determined in satisfying the conditions $\varphi^\alpha(y(t_0), y'(t_0)) = 0$, since the matrix (1:3) has rank m. That this solution in reality only depends upon 2n-2 arbitrary constants instead of 2n, is a consequence of the problem being in parametric form.

5. EXTREMALOIDS, NORMAL EXTREMALOIDS AND CONSTRUCTION OF FAMILIES OF EXTREMALOIDS.

An admissible arc and set of multipliers

$$
y^i = y^i(t), \quad \lambda^0, \quad \lambda^\alpha = \lambda^\alpha(t), \quad (t_0 \le t \le T),
$$

is called an extremaloid if the functions $y^i(t)$ are continuous and satisfy the differential equations (2:1) as well as the equations (2:2) at the corners. The functions $\lambda^\alpha(t)$ are continuous except possibly at the corners.

An extremaloid is said to be normal if there exists for it a set of 2n admissible variations for which the determinant

$$
(5:1) \qquad \begin{vmatrix} \eta^{ik}(t_0) \\ \eta^{ik}(T) \end{vmatrix} \qquad (k = 1, \ldots, 2n)
$$

is different from zero. It is normal on a sub-interval

(ξ_1, ξ_2) of (t_0, T) if there exists a set of 2n admissible variations for which the determinant (5:1) is not zero when t_0 and T are replaced by ξ_1, and ξ_2 respectively. Concerning normal extremaloids we have the following useful theorem.

THEOREM 5:1. _A necessary and sufficient condition that an extremaloid be normal is that there exists for it no set of multipliers_ λ^0, $\lambda^\alpha(t)$ _with_ $\lambda^0 = 0$. _For such a normal extremaloid multipliers in the form_ $\lambda^0 = 1$, $\lambda^\alpha(t)$ _always exist and in this form they are unique._

This has been proved by Bliss[6] for extremals in non-parametric form. However the proof he gives is valid for extremaloids in parametric form when we change his x to t.

We wish next to construct a family of extremaloids containing a given extremaloid. Suppose that E_0 is an extremaloid joining two points y_0, and Y, having corners at $y_1, \ldots, y_p$, and suppose that $R \neq 0$ along E_0 including both sides of the corners. Let

$$y^i = \psi^i(t, b), \qquad \lambda^\alpha = \Lambda^\alpha(t, b),$$

be a family of extremals containing the first extremal arc E_{01}, of E_0 for $b = b_0$, $t_0 \leq t \leq t_1$. Let p^{i-} and p^{i+} represent the direction cosines of the tangents of E_0 at the corner y_1, and for convenience suppose $\psi^{i\prime}(t_1, b_0) = p^{i-}$. Let $\Gamma^{\alpha+}$ and $\Gamma^{\alpha-}$ be the value of the multipliers of E_0 at the corner y_1, so that $\Lambda^\alpha(t_1, b_0) = \Gamma^{\alpha-}$. The equations

6. _The problem of Lagrange in the calculus of variations_, American Journal of Mathematics, vol. 52(1930) p. 688.

$$
\begin{aligned}
&F_{y^{i\prime}}[\psi(t,b),q,M] - F_{y^{i\prime}}[\psi(t,b),\psi'(t,b),\Lambda(t,b)] = 0, \\
(5:2) \qquad &\varphi^{\beta}[\psi(t,b),q] = 0, \\
&q^{j}q^{j} - 1 = 0,
\end{aligned}
$$

have the initial solution $b = b_o$, $t = t_1$, $q^{i} = p^{i+}$, $M^{\alpha} = \Gamma^{\alpha+}$. The functional determinant of equations (5:2) with respect to q, M and t, is

$$
\Delta = \begin{vmatrix} F^{+}_{y^{i\prime}y^{j\prime}} & F^{+}_{y^{i\prime}\lambda^{\alpha}} & [F^{+}_{y^{i\prime}y^{j}}p^{j-} - F^{-}_{y^{i}}] \\ \varphi^{\beta+}_{y^{j\prime}} & 0 & \varphi^{\beta+}_{y^{j}}p^{j-} \\ 2p^{j+} & 0 & 0 \end{vmatrix},
$$

since E_o satisfies the Euler-Lagrange equations (2:1). If we expand this determinant by the last row and column, making use of theorem 3:1 and the identities (3:1) we find that the determinant reduces to $-2F_1^{+}\Omega_o$ where

$$\Omega_o \equiv p^{i-}F^{+}_{y^{i}} - p^{i+}F^{-}_{y^{i}}.$$

Hence if we make the additional assumption that $\Omega_o \neq 0$ we have $\Delta \neq 0$ and equations (5:2) have unique solutions $t = t(b)$, $q^{i} = q^{i}(b)$, $M^{\alpha} = M^{\alpha}(b)$, having b, t, q and M in sufficiently restricted neighborhoods of b_o, t_1, p^{+} and Γ^{+} respectively, and these solutions are of class C'' at least. Application of the imbedding theorem 4:1 to the extremal arc E_{12}, adjacent to E_{01} shows that through each point $y^{i} = \psi^{i}(t(b),b)$ there passes a unique extremal arc in the admissible directions $y^{i\prime} = q^{i}(b)$ with multipliers $M^{\alpha} = \Lambda^{\alpha}(t(b),b)$, provided b is sufficiently close to b_o. This leads to the following theorem.

THEOREM 5:2. _If the function Ω_o is different from zero at each of the corners on E_o, we obtain by the above method_ (with proper choice of the parameter t on each sub-arc)

<u>a family</u> $y^i = y^i(t,b)$, $\lambda^\alpha = \lambda^\alpha(t,b)$ <u>of extremaloids defined for</u> $t_o - \delta \leq t \leq T + \delta$ <u>containing the original extremaloid</u> E_o <u>for</u> $b = b_o$, $t_o \leq t \leq T$. <u>The functions</u> $y^i, y^{i\prime}, y^{i\prime\prime}, \lambda^\alpha$ <u>are of class</u> C'' <u>in all their arguments except at corners.</u>

Suppose $y^i = \psi^i(t,b)$, $\lambda^\alpha = \Lambda^\alpha(t,b)$ is a 2n-2 parameter family of extremaloids constructed by the method given above. Later when we come to the determination of the conjugate points we shall be interested in the determinant

$$(5{:}3) \qquad \Theta = \begin{vmatrix} \dfrac{\partial\psi^i(t,b)}{\partial b^h} & \dfrac{\partial\psi^i(t,b)}{\partial t} & 0 \\ \dfrac{\partial\psi^i(t_o,b)}{\partial b^h} & 0 & \dfrac{\partial\psi^i(t_o,b)}{\partial t} \end{vmatrix},$$

both as to its zeros and its sign. In studying the determinant Θ, it is convenient to have the corners on all the extremaloids of the family occur for the same values of the parameter t. This can be brought about by a transformation[7] of the form $t = f(\bar{t},b)$ where f is continuous in all its arguments, f and f' are of class C' between corners and $f' > 0$. Such a transformation multiplies the determinant Θ by $f'(\bar{t})f'(\bar{t}_o)$ and hence cannot affect the sign nor the zeros of Θ. When the corners on all the extremaloids of the family occur for the same values of the parameter t, the partial derivatives of the functions ψ^i with respect to the parameters b^h are continuous functions of t.

---

7. Graves, L.M., <u>Discontinuous solutions in space problem of the calculus of variations</u>, American Journal of Mathematics, vol. 52(1930), p. 7.

6. A RELATION BETWEEN THE FUNCTIONS $\mathcal{E}$ AND Ω_o, AND A FURTHER NECESSARY CONDITION.

The Weierstrass necessary condition for a minimum[8] states that if E_o is an extremaloid which is normal on every sub-interval, then the Weierstrass function

$$\mathcal{E}(y,y',\bar{y}',\lambda) = \bar{y}^{i\prime}\left[F_{y^{i\prime}}(y,\bar{y}',\lambda) - F_{y^{i\prime}}(y,y',\lambda)\right] \geqq 0$$

for every (y,y',λ) on E_o and every $\bar{y}'$ satisfying $\varphi^\alpha(y,\bar{y}') = 0$. If t_k defines a corner of E_o and $y^{i\prime}(t_k^-) = p^{i-}$; $y^{i\prime}(t_k^+) = p^{i+}$, then (y,p^-,λ^-) and (y,p^+,λ^+) are each elements of E_o at the point $t = t_k$ and (y,p^+) and (y,p^-) satisfy the equations (1:1). Hence at the point $t = t_k$ we have two admissible sets (y,p^-,p^+,λ^-) and (y,p^+,p^-,λ^+). Making use of the corner equations (2:2) we find

$$\mathcal{E}(y,p^-,p^+,\lambda^-) = \varphi^\alpha(y,p^+)(\lambda^{\alpha-} - \lambda^{\alpha+}) = 0. \tag{6:1}$$

And in the same manner we have

$$\mathcal{E}(y,p^+,p^-,\lambda^+) = 0. \tag{6:2}$$

Suppose we select the arc length as the parameter along the extremaloid E_o and let (y,y',λ) define that branch of E_o preceding a corner defined for the parameter value s_k. At the same time let $\bar{y}^{i\prime}(s)$ be a set of functions which satisfy the equations $\varphi^\alpha(y,\bar{y}') = 0$ and have $\bar{y}^{i\prime}(s_k) = p^{i+}$. If we substitute these functions in the Weierstrass function, differentiate with respect to s, set $s = s_k$, and make use of the identities (3:1), $F_{y^{i\prime}\lambda^\alpha} = \varphi^\alpha_{y^{i\prime}}$ and the corner equations (2:2), we find

---

8. Eshleman, Dissertation, p. 12.

$$\frac{d}{ds}\mathcal{E}(y,p^-,p^+,\lambda^-) = \bar{y}^{i\prime\prime}\varphi^\alpha_{y^{i\prime}}(y,p^+)(\lambda^{\alpha-}-\lambda^{\alpha+})$$
$$+p^{j-}F_{y^j}(y,p^+,\lambda^-) - p^{i+}F_{y^i}(y,p^-,\lambda^-).$$

By differentiating the equations $\varphi^\alpha(y,\bar{y}') = 0$ with respect to s and setting $s = s_k$ we arrive at the equations

$$\varphi^\alpha_{y^{i\prime}}(y,p^+)\bar{y}^{i\prime\prime} = -\varphi^\alpha_{y^i}(y,p^+)p^{i-}.$$

Placing these in the above we find

$$\frac{d}{ds}\mathcal{E}(y,p^-,p^+,\lambda^-) = \Omega_o. \tag{6:3}$$

Suppose now we let the functions (y,y',λ) define that branch of E_o just past the corner defined by $s = s_k$, while at the same time let $\bar{y}^{i\prime}(s)$ satisfy the equations $\varphi^\alpha(y,\bar{y}') = 0$ and having $\bar{y}^{i\prime}(s_k) = p^{i-}$. Then if we repeat the above process we find that

$$\frac{d}{ds}\mathcal{E}(y,p^+,p^-,\lambda^+) = -\Omega_o. \tag{6:4}$$

From the Weierstrass condition and the equations (6:1), (6:2), (6:3) and (6:4) we readily obtain the following theorems.

THEOREM 6:1. *If E_o is an extremaloid along which $\mathcal{E} \geq 0$, then $\Omega_o \leq 0$ at the corners of E_o.*

THEOREM 6:2. *If E_o is an admissible arc normal on every sub-interval and minimizes the integral I then $\Omega_o \leq 0$ at the corners of E_o.*

7. THE EXTENSION OF THE ACCESSORY EQUATIONS AND PROPERTIES OF THEIR SOLUTIONS.

If the extremaloid E_o minimizes the integral I then the second variation[9]

---

9. Bliss, *The problem of Lagrange in the calculus of variations* American Journal of Mathematics, vol. 52 (1930).

$$J(\eta) = \int_{t_o}^{T} 2\omega(t,\eta,\eta')dt$$

is greater than or equal to zero for all admissible variations $\eta^i(t)$ such that $\eta^i(t_o) = \eta^i(T) = 0$, where the quadratic form ω is defined by the equation

$$2\omega(t,\eta,\eta') = \eta^i\eta^j F_{y^iy^j} + 2\eta^i\eta^{j'}F_{y^iy^{j'}} + \eta^{i'}\eta^{j'}F_{y^{i'}y^{j'}}.$$

Here and in the remaining sections it is to be understood that the arguments of F are the functions $y^i(t)$, $\lambda^\alpha(t)$ defining the extremaloid E_o, and it is assumed throughout that E_o is an extremaloid normal on every sub-interval, that $R \neq 0$ along E_o, and that $\Omega_o \neq 0$ at the corners.

Since admissible variations $\eta^i(t)$ must be of class D' on (t_o, T) and satisfy the differential equations

(7:1) $$\Phi^\alpha(\eta,\eta') = \varphi^\alpha_{y^i}\eta^i + \varphi^\alpha_{y^{i'}}\eta^{i'} = 0,$$

it is clear that the properties of the second variation suggest a minimum problem in $t\eta$-space of the same type as the original Lagrange problem in ty-space. Hence to the second variation we can apply the necessary conditions of section 2 which are valid even though the integrand function is discontinuous in t. If we let

(7:2) $$\Omega \equiv \omega + \mu^\alpha\Phi^\alpha$$

then the multiplier rule for the minimum problem in $t\eta$-space may be stated as follows

THEOREM 7:1. For every minimizing arc $\eta^i = \eta^i(t)$, $(t_o \leq t \leq T)$, giving to the second variation J its minimum value zero, there exists a set of constants C^i and a set of functions $\mu^\alpha(t)$ of class D such that the equations

$$\Omega_{\eta^{i\prime}} = \int_{t_0}^{t} \Omega_{\eta^i}\, dt + c^i, \qquad \Phi^\alpha = 0 \tag{7:3}$$

<u>hold along the minimizing arc.</u>

It is to be noted that since equations (7:1) are linear in $\eta^i, \eta^{i\prime}$ they are their own equations of variations for the problem in $t\eta$-space. Hence when E_0 is a normal extremaloid for the original problem, every extremaloid for the problem in $t\eta$-space will be normal and by theorem (5:1) we are justified in taking $\mu^0 = 1$ in equations (7:2).

The equations (7:3) are to be looked upon as the <u>extension of the accessory equations</u> for the case of discontinuous solutions. They are equivalent to the differential equations

$$\frac{d}{dt}\Omega_{\eta^{i\prime}} = \Omega_{\eta^i}, \qquad \Phi^\alpha = 0, \tag{7:4}$$

holding between corners of E_0 and η^i, and the corner equations

$$\Omega_{\eta^{i\prime}}(t^-, \eta, \eta'^-, \mu^-) = \Omega_{\eta^{i\prime}}(t^+, \eta, \eta'^+, \mu^+). \tag{7:5}$$

LEMMA 7:1. <u>Every set of functions of the form</u> $\eta^i = \rho y^{i\prime}$, $\mu^\alpha = \rho\lambda^{\alpha\prime}$ <u>is a solution of equations</u> (7:3) <u>provided</u> ρ <u>is a function of</u> t <u>of class</u> D' <u>on</u> (t_0, T) <u>and</u> $\rho(t_1) = ,\dots, = \rho(t_p) = 0$, <u>where</u> $t_1, \dots, t_p$, <u>define the corner points on</u> E_0.

In order to prove this lemma we note that

$$\begin{aligned} \Omega_{\eta^{i\prime}}(t, \rho y', \rho y'' + \rho' y', \rho\lambda') &= \rho y^{j\prime} F_{y^{i\prime} y^j} + \rho y^{j\prime\prime} F_{y^{i\prime} y^{j\prime}} \\ &\quad + \rho' y^{j\prime} F_{y^{i\prime} y^{j\prime}} + \rho\lambda^{\alpha\prime}\varphi^\alpha_{y^{i\prime}} \\ &= \rho \frac{d}{dt}(F_{y^{i\prime}}) = \rho F_{y^i}, \end{aligned} \tag{7:6}$$

on account of the homogeneity of F and the Euler-Lagrange

equations (2:1). Similarly

$$
(7:7)\qquad \begin{aligned}\Omega_{\eta^{i}}(t,\rho y',\rho y''+\rho' y',\rho\lambda') &= \rho y^{j\prime} F_{y^{i} y^{j}} + \rho y^{j\prime\prime} F_{y^{i} y^{j\prime}} \\ &\quad + \rho' y^{j\prime} F_{y^{i} y^{j\prime}} + \rho \lambda^{\alpha\prime} \varphi^{\alpha}_{y^{i}} \\ &= \rho \frac{d}{dt}(F_{y^{i}}) + \rho' F_{y^{i}} = \frac{d}{dt}(\rho F_{y^{i}}).\end{aligned}
$$

Hence we have

$$\frac{d}{dt}\Omega_{\eta^{i\prime}}(t,\rho y',\rho y''+\rho' y',\rho\lambda') = \Omega_{\eta^{i}}(t,\rho y',\rho y''+\rho' y',\rho\lambda').$$

The second set of equations (7:4) are also seen to be satisfied for we have

$$
\begin{aligned}\Phi^{\alpha}(t,\rho y',\rho y''+\rho' y') &= \rho y^{i\prime}\varphi^{\alpha}_{y^{i}} + \rho y^{i\prime\prime}\varphi^{\alpha}_{y^{i\prime}} + \rho' y^{i\prime}\varphi^{\alpha}_{y^{i\prime}} \\ &= \rho\frac{d}{dt}(\varphi^{\alpha}) + \rho'\varphi^{\alpha} = 0,\end{aligned}
$$

since the functions $y^{i}(t)$ defining E_0 are admissible. Hence the complete set of equations (7:4) is satisfied for these values. Moreover the corner equations (7:5) are satisfied since ρ is zero at the corner points of E_0 and ρ' disappears from the final form of $\Omega_{\eta^{i\prime}}$ in (7:6).

We shall say that a solution $(\eta(t),\mu(t))$ of equations (7:4) is <u>orthogonal</u> on an interval (t_{k-1},t_k) in case the functions $\eta^{i}(t)$ are of class C' and satisfy the equation $y^{i\prime}\eta^{i\prime} = 0$, where the functions $y^{i\prime}(t)$ are those belonging to E_0, and the functions μ^{α} are continuous on the interval. Using the assumption that $R \neq 0$ along E_0 it is easily proved that for every orthogonal solution the functions η^{i} are of class C'', and the functions μ^{α} are of class C'.[10]

10. Bliss, <u>Jacobi's condition for problems of the calculus of variations in parametric form</u>, Trans. Amer. Math. Soc. vol. 17(1916) pp. 195-206.

LEMMA 7:2. *Let* (η,μ) *be an orthogonal solution of equations* (7:4) *such that*

$$\eta^i(\tau) = \eta^{i\prime}(\tau) = \mu^\alpha(\tau) = 0$$

for some value τ *in the interval* (t_{k-1}, t_k). *Then*

$$\eta^i(t) \equiv \eta^{i\prime}(t) \equiv \mu^\alpha(t) \equiv 0$$

on $t_{k-1} \leq t \leq t_k$.

The proof of this lemma as given by Eshleman[11] needs but a slight modification due to our difference in defining orthogonality.

LEMMA 7:3. *To each solution* (η,μ) *of equations* (7:4) *having* η^i *of class* C' *and* μ^α *continuous there corresponds a unique orthogonal solution of the form* $(\eta - \rho y', \mu - \rho\lambda')$, *where* ρ *is a function of* t *of class* C' *having a prescribed value at a single point.*

That $\eta_0^i \equiv \eta^i - \rho y^{i\prime}$, $\mu_0^\alpha \equiv \mu^\alpha - \rho\lambda^{\alpha\prime}$, is a solution of equations (7:4) is at once obvious from lemma 7:1 . The condition for orthogonality gives

$$y^{i\prime}\eta_0^{i\prime} = y^{i\prime}(\eta^{i\prime} - \rho' y^{i\prime} - \rho y^{i\prime\prime}) = y^{i\prime}\eta^{i\prime} - \rho' y^{i\prime} y^{i\prime} - \rho y^{i\prime} y^{i\prime\prime} = 0.$$

This is a linear differential equation of the first order in ρ, hence ρ is uniquely determined by prescribing its value at a single point.

LEMMA 7:4. *If* (η,μ) *is a solution of equations* (7:3) *having*

$$\eta^i(\tau) = \gamma y^{i\prime}(\tau), \quad \eta^{i\prime}(\tau) = \gamma y^{i\prime\prime}(\tau) + \delta y^{i\prime}(\tau), \tag{7:8}$$
$$\mu^\alpha(\tau) = \lambda^{\alpha\prime}(\tau),$$

---

11. Dissertation, p. 21.

<u>at a single point</u> τ, <u>then</u> $\eta^i = \rho y^{i\prime}$ <u>and</u> $\mu^\alpha = \rho\lambda^{\alpha\prime}$ <u>on the whole interval</u> (t_0, T), <u>where</u> ρ <u>is a function of class</u> D' <u>and vanishes at the parameter values</u> $t_1, \ldots, t_p$, <u>defining corners of</u> E_0.

Consider an interval containing τ on which the η^i are of class C', and the μ^α are continuous. Consider the associated orthogonal solution

$$\eta_0^i = \eta^i - \rho y^{i\prime}, \quad \mu_0 = \mu^\alpha - \rho\lambda^{\alpha\prime}, \tag{7:9}$$

having $\rho(\tau) = \gamma$. By lemma (7:3) this orthogonal solution is uniquely determined. Then

$$\eta_0^i(\tau) = \eta^i(\tau) - \gamma y^{i\prime}(\tau) = 0.$$

The condition of orthogonality at the point τ gives

$$\begin{aligned} 0 &= y^{i\prime}(\tau)\left[\eta^{i\prime}(\tau) - \rho'(\tau)y^{i\prime}(\tau) - \gamma y^{i\prime\prime}(\tau)\right] \\ &= y^{i\prime}(\tau)y^{i\prime}(\tau)\left[\delta - \rho'(\tau)\right]. \end{aligned}$$

Hence, $\delta = \rho'(\tau)$.

Using this value of $\rho'(\tau)$ we find

$$\begin{aligned} \eta_0^{i\prime}(\tau) &= \eta^{i\prime}(\tau) - \delta y^{i\prime}(\tau) - \gamma y^{i\prime\prime}(\tau) \\ &= \gamma y^{i\prime\prime}(\tau) + \delta y^{i\prime}(\tau) - \delta y^{i\prime}(\tau) - \gamma y^{i\prime\prime}(\tau) = 0. \end{aligned}$$

Considering $\mu_0^\alpha(\tau)$, we have from (7:9) and (7:8)

$$\mu_0^\alpha(\tau) = \mu^\alpha(\tau) - \rho(\tau)\lambda^{\alpha\prime}(\tau) = 0.$$

Hence we have

$$\eta_0^i(\tau) = \eta_0^{i\prime}(\tau) = \mu_0^\alpha(\tau) = 0.$$

This is the hypothesis of lemma 7:2 and it follows at once that $\eta_0^i(t) \equiv 0, \mu_0^\alpha(t) \equiv 0$ on the interval on which η^i is of class C' and μ^α is continuous.

Next, let τ be a point of discontinuity of $\eta^{i\prime}$ different from $t_1, \ldots, t_p$, the parameter values defining the

corners of E_o. We consider the corner equations (7:5) at this point and find

$$\Omega^{+}_{\eta^{i\prime}} = F_{y^j y^{i\prime}}\eta^j + F_{y^{i\prime} y^{j\prime}}\eta^{j\prime +} + \mu^{\alpha +}\varphi^{\alpha}_{y^{i\prime}},$$
$$\Omega^{-}_{\eta^{i\prime}} = F_{y^j y^{i\prime}}\eta^j + F_{y^{i\prime} y^{j\prime}}\eta^{j\prime -} + \mu^{\alpha -}\varphi^{\alpha}_{y^{i\prime}},$$

from which we see at once that the corner equations reduce to

$$F_{y^{i\prime} y^{j\prime}}(\eta^{j\prime +} - \eta^{j\prime -}) + \varphi^{\alpha}_{y^{i\prime}}(\mu^{\alpha +} - \mu^{\alpha -}) = 0.$$

We have given that (η, μ) is a solution of equations (7:3), hence $\Phi^{\alpha}(\eta, \eta') = 0$, at both sides of the corners and we have

$$\varphi^{\alpha}_{y^i}\eta^i + \varphi^{\alpha}_{y^{i\prime}}\eta^{i\prime +} = 0,$$
$$\varphi^{\alpha}_{y^i}\eta^i + \varphi^{\alpha}_{y^{i\prime}}\eta^{i\prime -} = 0.$$

Subtracting the latter from the former we have

$$\varphi^{\alpha}_{y^{i\prime}}(\eta^{i\prime +} - \eta^{i\prime -}) = 0.$$

This leads to the system of equations

$$F_{y^{i\prime} y^{j\prime}}(\eta^{j\prime +} - \eta^{j\prime -}) + \varphi^{\alpha}_{y^{i\prime}}(\mu^{\alpha +} - \mu^{\alpha -}) = 0,$$
$$\varphi^{\beta}_{y^{j\prime}}(\eta^{j\prime +} - \eta^{j\prime -}) = 0,$$

which is linear in the variables $(\eta^{j\prime +} - \eta^{j\prime -}), (\mu^{\alpha +} - \mu^{\alpha -})$ and the determinant of their coefficients is D. Such a system of equations is known to have the solution $(y', 0)$ which is not identically zero, and as D has rank $n + m - 1$ we conclude at once that

$$\mu^{\alpha +} - \mu^{\alpha -} = 0$$
$$\eta^{j\prime +} - \eta^{j\prime -} = k y^{j\prime}.$$

Hence we have on the forward side of the corner

$$\eta^{i+}(\tau) = \eta^{i-}(\tau) = \gamma y^{i\prime}(\tau),$$
$$\eta^{i\prime +}(\tau) = \eta^{i\prime -}(\tau) + k y^{i\prime}(\tau)$$
$$= \gamma y^{i\prime\prime}(\tau) + (\delta + k) y^{i\prime}(\tau),$$

$$\mu^{\alpha+}(\tau) = \mu^{\alpha-}(\tau) = \lambda^{\alpha\prime}(\tau).$$

Thus the hypothesis of the lemma holds on the opposite side of the corner and we may apply that part of the lemma which has already been proved. Thus the solution (η, μ) has the form $(\rho y', \rho\lambda')$ on the adjacent interval on which η^i is of class C', and μ^α is continuous.

Finally, suppose the hypothesis of the lemma holds at $\tau = t_k^-$ corresponding to one side of one of the corners on E_0. Considering the corner equations (7:5) at this point and making use of the Euler-Lagrange equations and the identities (3:1) we find

$$\Omega^-_{\eta^{i\prime}} = F^-_{y^{i\prime}y^j}\gamma y^{j\prime-} + F^-_{y^{i\prime}y^{j\prime}}[\gamma y^{j\prime\prime-} + \delta y^{j\prime-}] + \gamma\lambda^{\alpha\prime-}\varphi^{\alpha-}_{y^{i\prime}}$$
$$= \gamma\tfrac{d}{dt}(F^-_{y^{i\prime}}) = \gamma F^-_{y^i},$$
$$\Omega^+_{\eta^{i\prime}} = F^+_{y^{i\prime}y^j}\gamma y^{j\prime-} + F^+_{y^{i\prime}y^{j\prime}}\eta^{j\prime+} + \mu^{\alpha+}\varphi^{\alpha+}_{y^{i\prime}}.$$

Hence the corner equations (7:5) yield

$$\gamma F^-_{y^i} - F^+_{y^{i\prime}y^{j\prime}}\eta^{j\prime+} - \gamma F^+_{y^{i\prime}y^j}y^{j\prime-} - \mu^{\alpha+}\varphi^{\alpha+}_{y^{i\prime}} = 0. \tag{7:10}$$

If we multiply equations (7:10) by $y^{i\prime+}$ and add we find

$$0 = \gamma y^{i\prime+}F^-_{y^i} - y^{i\prime+}F^+_{y^{i\prime}y^{j\prime}}\eta^{j\prime+} - \gamma y^{i\prime+}F^+_{y^{i\prime}y^j}y^{j\prime-}$$
$$- \mu^{\alpha+}y^{i\prime+}\varphi^{\alpha+}_{y^{i\prime}}$$
$$= \gamma y^{i\prime+}F^-_{y^i} - \gamma y^{j\prime-}F^+_{y^j}$$
$$= -\gamma\Omega_0$$

Since by hypothesis $\Omega_0 \neq 0$ we have $\gamma = 0$. Placing this value in equations (7:10) we have

$$F^+_{y^{i\prime}y^{j\prime}}\eta^{j\prime+} + \mu^{\alpha+}\varphi^{\alpha+}_{y^{i\prime}} = 0.$$

Again from $\Phi^\alpha = 0$ we have $\varphi^{\beta+}_{y^j}\eta^{j+} + \varphi^{\beta+}_{y^{j\prime}}\eta^{j\prime+} = 0$, which reduces to $\varphi^{\beta+}_{y^{j\prime}}\eta^{j\prime+} = 0$, since $\gamma = 0$.

Hence we have the system of equations

$$F^{+}_{y^{i\prime} y^{j\prime}} \eta^{j\prime +} + \mu^{\alpha +} \varphi^{\alpha +}_{y^{i\prime}} = 0,$$
$$\varphi^{\beta +}_{y^{j\prime}} \eta^{j\prime +} \qquad\qquad = 0,$$

which is linear in the variables $\eta^{j\prime +}$, $\mu^{\alpha +}$, and their determinant of coefficients is D. As we argued before we find $\mu^{\alpha +} = 0$, $\eta^{j\prime +} = ky^{i\prime}$, and we have that the hypothesis of our lemma holds on the opposite side of the corner.

8. DEFINITION OF CONJUGATE POINT AND THE EXTENSION OF THE MAYER CONDITION.

A point y is defined to be conjugate to a point y_0 on an extremaloid E_0 in case it corresponds to a parameter value $\tau \neq t_0$ such that there exist constants $\varepsilon, \gamma, \delta$, and an admissible solution (η, μ) of the accessory equations (7:3) with the properties

1) (η, μ) is not of the form $(\rho y', \rho \lambda')$;
2) $\varepsilon \eta^{i}(t_0) = 0$, and $\varepsilon \eta^{i}(\tau) = \gamma y^{i\prime}(\tau^-) + \delta y^{i\prime}(\tau^+)$;
3) $\varepsilon, \gamma, \delta$, are not all zero and $\gamma\delta \geqslant 0$;

where $y^{i\prime}(\tau^-)$ and $y^{i\prime}(\tau^+)$ refer to the left and right hand derivatives of $y^{i}(t)$ at the point $t = \tau$, respectively. The special case where γ and δ may both be taken as zero is the only case that is of interest for a minimizing arc without corners. In case $\varepsilon = 0$ it is obvious since $y^{i\prime} y^{i\prime} \neq 0$ that $\gamma\delta > 0$ and τ is at a corner which is a cusp on E_0.

THEOREM 8:1. EXTENSION OF THE MAYER CONDITION. *If E_0 is an admissible arc, normal on every sub-interval, joining the two points y_0 and Y and minimizing the integral J, and if $R \neq 0$ on E_0 and $\Omega_0 \neq 0$ at the corners on E_0, then there can be no point y conjugate to y_0 on E_0 between y_0 and Y.*

COROLLARY. <u>A minimizing arc can contain no cusp.</u>

The hypotheses imply that E_o is an extremaloid whose extremal arcs are each of class C''. Moreover $\Omega_o < 0$ at the corners by section 6. Suppose $t = \tau$ defines a point y that is conjugate to y_o and lies between y_o and Y, and let (η, μ) be a solution of the accessory equations defining the conjugate point y. Let

$$(8:1) \qquad \begin{aligned} \xi^i &\equiv \varepsilon\eta^i - \rho y^{i\prime}, & u^\alpha &\equiv \varepsilon\mu^\alpha - \rho\lambda^{\alpha\prime} & &\text{on } (t_o \leq t \leq \tau) \\ &\equiv \rho y^{i\prime}, & &\equiv \rho\lambda^{\alpha\prime} & &\text{on } (\tau \leq t \leq T), \end{aligned}$$

where the function ρ vanishes at $t_o, t_1, \ldots, t_p, T$, except at $t = \tau$ where $\rho(\tau^-) = \gamma$, $\rho(\tau^+) = \delta$. If we assume that ρ is of class D' on (t_o, τ) and on (τ, T) then ξ^i is an admissible variation as it is continuous at $t = \tau$ by the second part of property 2), and ξ^i satisfies equations $\Phi^\alpha = 0$ since η^i does by hypothesis and lemma 7:1 gives $\rho y^{i\prime}$ as a solution. The functions ξ^i also vanish at t_o and T and with u^α satisfy the accessory equations (7:3) on the intervals (t_o, τ) and (τ, T). However, (ξ, u) cannot satisfy the accessory equations (7:3) on the whole interval (t_o, T) since by lemma (7:4) if $\varepsilon \neq 0$, (ξ, u) has the form $(\rho y', \rho\lambda')$ only on part of the interval, and if $\varepsilon = 0$ then $\gamma\delta > 0$, and the parameter value τ corresponds to a corner at which ρ does not vanish. Consequently $J(\xi)$ cannot be a minimum. We proceed to show that $J(\xi) \leq 0$ and hence E_o cannot minimize the integral I if E_o contains a point y conjugate to y_o as assumed.

We wish to consider the integral

$$J(\xi) = \int_{t_0}^{T} 2\omega(\xi, \xi')dt.$$

Since the equations $\Phi^\alpha = 0$ are satisfied we can replace our integrand function 2ω by 2Ω and we have

$$J(\xi) = \int_{t_0}^{T} 2\Omega(\xi, \xi', u)dt.$$

Let us consider the integral on an interval (t_{k-1}, t_k) that does not contain τ. If we make use of the well-known property of quadratic forms

$$2\Omega = \xi^i \Omega_{\xi^i} + \xi^{i\prime}\Omega_{\xi^{i\prime}} + u^\alpha \Omega_{u^\alpha} \,,$$

equations (7:4), and the equations $\Omega_{u^\alpha} = \Phi^\alpha = 0$, we find

$$\begin{aligned}\int_{t_{k-1}}^{t_k} 2\Omega\, dt &= \int_{t_{k-1}}^{t_k} (\xi^{i\prime}\Omega_{\xi^{i\prime}} + \xi^i \frac{d}{dt}\Omega_{\xi^{i\prime}})dt \\ &= \xi^i \Omega_{\xi^{i\prime}}(\xi, \xi', u)\Big|_{t_{k-1}}^{t_k} \,.\end{aligned}$$

Since ξ^i vanishes at t_0 and T and since $\Omega_{\xi^{i\prime}}$ is continuous we have

$$J(\xi) = \int_{t_0}^{T} 2\Omega dt = \xi^i \Omega_{\xi^{i\prime}}(\xi^i, \xi^{i\prime}, u)\Big|_{\tau^+}^{\tau^-} \,.$$

Making use of the definitions (8:1) and equations (7:6) we find

$$\begin{aligned}J(\xi) &= \xi^i\left[\epsilon\, \Omega^-_{\eta^{i\prime}}(\eta, \eta', \mu) - \Omega^-_{\eta^{i\prime}}(\rho y', \rho y'' + \rho' y', \rho\lambda')\right] \\ &\quad - \xi^i \Omega^+_{\eta^{i\prime}}(\rho y', \rho y'' + \rho' y', \rho\lambda') \\ &= \xi^i\left[\epsilon\, \Omega^-_{\eta^{i\prime}}(\eta, \eta', \mu) - \gamma F^-_{y^{i\prime}}\right] - \xi^i F^+_{y^{i\prime}}\, \delta \,.\end{aligned}$$

As we have noted above ξ^i is continuous at the point $t = \tau$. Hence in the value of $J(\xi)$ above we replace the first ξ^i by its value on the right of τ and the second by its value on the

left and we have

$$J(\xi) = \delta y^{i\prime +}\left[\varepsilon \Omega^{-}_{\eta^{i\prime}}(\eta,\eta',\mu) - \gamma F^{-}_{y^{i}}\right] - (\varepsilon\eta^{i} - \gamma y^{i\prime -})\delta F^{+}_{y^{i}}$$
$$= \varepsilon\delta y^{i\prime +}\Omega^{+}_{\eta^{i\prime}}(\eta,\eta',\mu) - \gamma\delta(y^{i\prime +}F^{-}_{y^{i}} - y^{i\prime -}F^{+}_{y^{i}})$$
$$- \varepsilon\delta\eta^{i}F^{+}_{y^{i}},$$

since $\Omega_{\eta^{i\prime}}$ is continuous on the solution (η,μ). Making use of the identity

$$y^{i\prime +}\Omega^{+}_{\eta^{i\prime}}(\eta,\eta',\mu) = y^{i\prime +}(F^{+}_{y^{j}y^{i\prime}}\eta^{j} + F^{+}_{y^{j\prime}y^{i\prime}}\eta^{j\prime +}$$
$$+ \mu^{\alpha +}\varphi^{\alpha +}_{y^{i\prime}})$$
$$= F^{+}_{y^{j}}\eta^{j} + \mu^{\alpha +}y^{i\prime +}\varphi^{\alpha +}_{y^{i\prime}} = F^{+}_{y^{j}}\eta^{j},$$

we find

$$J(\xi) = \varepsilon\delta F^{+}_{y^{j}}\eta^{j} - \gamma\delta(y^{i\prime +}F^{-}_{y^{i}} - y^{i\prime -}F^{+}_{y^{i}}) - \varepsilon\delta F^{+}_{y^{i}}\eta^{i} = \gamma\delta\Omega_{0}.$$

Hence, $J(\xi) = 0$ if y is not at a point which is a corner of E_0 and $J(\xi) = \gamma\delta\Omega_0 \leq 0$ if y is at a corner of E_0.

9. DETERMINATION OF THE CONJUGATE POINTS.

Associated with a set of 2n+2 solutions (η^{h},μ^{h}) of the accessory equations (7:3) is an important determinant

$$\Theta(t,t_0) = \begin{vmatrix} \eta^{ih}(t) & y^{i\prime}(t) & 0 \\ \eta^{ih}(t_0) & 0 & y^{i\prime}(t_0) \end{vmatrix},$$

and a matrix

$$M = \begin{Vmatrix} \eta^{ih} & y^{i\prime} & 0 \\ \eta^{ih\prime} & y^{i\prime\prime} & y^{i\prime} \\ \mu^{\alpha h} & \lambda^{\alpha\prime} & 0 \end{Vmatrix}.$$

We shall proceed to develop the relations between Θ, and M, as well as their relation to the points conjugate to y_0 and the family of extremaloids discussed in section 5.

LEMMA 9:1. _If_ (η^{h},μ^{h}) _are_ 2n-2 _solutions of the equations_ (7:3) _such that columns of the matrix_ M _are linearly_

independent at one point $t = \tau$, *then the columns of* M *are linearly independent for each value of* t *on* (t_o, T) *and every solution* (η, μ) *of equations* (7:3) *is expressible in the form*

$$\eta^i = A^h \eta^{ih} + \rho y^{i\prime}, \qquad \mu^\alpha = A^h \mu^{\alpha h} + \rho \lambda^{\alpha\prime}$$

where the A^h *are constants and* ρ *is a function of* t *of class* D' *vanishing at the parameter values* $t_1, \ldots, t_p$ *corresponding to corners on* E_o.

Assume that the columns of M are linearly dependent for $t = \tau$. Then there exists a set of constants A^h, γ, δ, not all zero satisfying the equations

$$\begin{aligned} A^h \eta^{ih} + \gamma y^{i\prime} &= 0, \\ A^h \eta^{ih\prime} + \gamma y^{i\prime\prime} + \delta y^{i\prime} &= 0, \\ A^h \mu^{\alpha h} + \gamma \lambda^{\alpha\prime} &= 0. \end{aligned}$$

Moreover the constants A^h are not all zero for if that were the case we would have $\gamma = \delta = 0$ also. The solution

$$\eta^i = A^h \eta^{ih}, \qquad \mu^\alpha = A^h \mu^{\alpha h}$$

of equations (7:3) has

$$\begin{aligned} \eta^i(\tau) &= -\gamma y^{i\prime}(\tau), \\ \eta^{i\prime}(\tau) &= -\gamma y^{i\prime\prime}(\tau) - \delta y^{i\prime}(\tau), \\ \mu^\alpha(\tau) &= -\gamma \lambda^{\alpha\prime}(\tau). \end{aligned}$$

Hence by lemma 7:4, (η, μ) has the form $(\rho y', \rho \lambda')$ on the whole interval (t_o, T) and thus the columns of M are linearly dependent throughout (t_o, T). Hence if the columns of M are linearly independent for one value of t they must be linearly independent throughout (t_o, T).

To complete the proof of the lemma, let (η, μ) be any solution of equations (7:3). Then at an arbitrary point $t = \tau$

of (t_0, T) the equations

$$
(9{:}1)\qquad
\begin{aligned}
\eta^i &= A^h\eta^{ih} + \gamma y^{i\prime},\\
\eta^{i\prime} &= A^h\eta^{ih\prime} + \gamma y^{i\prime\prime} + \delta y^{i\prime},\\
\mu^\alpha &= A^h\mu^{\alpha h} + \gamma\lambda^{\alpha\prime},
\end{aligned}
$$

are constants since (η, η'), $(\eta^h, \eta^{h\prime})$, (y', y''), $(0, y')$ all satisfy the equations $\Phi^\alpha = 0$ whose matrix of coefficients has rank m by hypothesis. Hence the equations (9:1) have a unique solution A^h, γ, δ. Then

$$\xi^i = \eta^i - A^h\eta^{ih}, \qquad \vartheta^\alpha = \mu^\alpha - A^h\mu^{\alpha h}$$

is a solution of equations (7:3) having

$$
\begin{aligned}
\xi^i(\tau) &= \gamma y^{i\prime}(\tau),\\
\xi^{i\prime}(\tau) &= \gamma y^{i\prime\prime}(\tau) + \delta y^{i\prime}(\tau),\\
\vartheta^\alpha(\tau) &= \gamma\lambda^{\alpha\prime}(\tau),
\end{aligned}
$$

and consequently (ξ, ϑ) has the form $(\rho y', \rho\lambda')$ on the whole interval (t_0, T) by lemma 7:4.

LEMMA 9:2. <u>If (η^h, μ^h) is a system of 2n-2 solutions of the equations (7:3) of class D' on (t_0, T) such that the matrix M has rank 2n, then the points y conjugate to y_0 on the extremaloid E_0 correspond to the parameter values $t = \tau \neq t_0$ at which $\Theta(t, t_0)$ vanishes or changes sign.</u>

We first show that Θ either vanishes or changes sign at a conjugate point. Suppose τ defines a point y conjugate to y_0. Then by section 8 there exists a solution (η, μ) of equations (7:3) with the properties:

1) (η, μ) is not of the form $(\rho y', \rho\lambda')$.
2) $\varepsilon\eta^i(t_0) = 0$; $\varepsilon\eta^i(\tau) = \gamma y^{i\prime}(\tau^-) + \delta y^{i\prime}(\tau^+)$.
3) $\varepsilon, \gamma, \delta$, are not all zero and $\gamma\delta \geqq 0$.

Since the columns of M are linearly independent (η, μ) is expressible in the form

$$\eta^i = A^h \eta^{ih} + \rho y^{i\prime}, \qquad \mu^\alpha = A^h \mu^{\alpha h} + \rho \lambda^{\alpha\prime}$$

where the constants A^h are not all zero and ρ vanishes at the parameter values $t_1, \ldots, t_p$ defining corners of E_o. Hence we have the system of linear equations

$$\begin{aligned} &\varepsilon A^h \eta^{ih}(\tau) + \varepsilon\rho y^{i\prime}(\tau) - \gamma y^{i\prime}(\tau^-) - \delta y^{i\prime}(\tau^+) = 0, \\ &\varepsilon A^h \eta^{ih}(t_o) + \varepsilon\rho y^{i\prime}(t_o) \qquad\qquad\qquad = 0. \end{aligned}$$

Consider first the case $\gamma\delta = 0$, in which we have $\varepsilon \neq 0$. Then obviously $\Theta(t, t_o) = 0$ for $t = \tau, \tau^-$ or τ^+. In case $\gamma\delta > 0$ we may suppose that y is at a corner, since otherwise the solution (η, μ) could be modified to make $\gamma\delta = 0$. Then $\rho(\tau) = 0$ and we must have

$$\begin{vmatrix} \eta^{ih}(\tau) & [\gamma y^{i\prime}(\tau^-) + \delta y^{i\prime}(\tau^+)] & 0 \\ \eta^{ih}(t_o) & 0 & y^{i\prime}(t_o) \end{vmatrix} = 0,$$

from which we see that Θ changes sign for $t = \tau$.

We next show that if Θ vanishes or changes sign we have a conjugate point. Suppose $\Theta(\tau, t_o) = 0$. Then the equations

$$\begin{aligned} A^h \eta^{ih}(\tau) + \gamma y^{i\prime}(\tau) &= 0, \\ A^h \eta^{ih}(t_o) + \delta y^{i\prime}(t_o) &= 0, \end{aligned}$$

have solutions A^h, γ, δ, not all zero. Moreover not all the A^h are zero for it that were the case γ and δ would also have to vanish since not all the $y^{i\prime}$ are zero. Let

$$\eta^i = A^h \eta^{ih} + \rho y^{i\prime}, \qquad \mu^\alpha = A^h \mu^{\alpha h} + \rho \lambda^{\alpha\prime},$$

where $\rho(t_o) = \delta$, $\rho(\tau) = \rho(t_1) = \ldots, = \rho(t_p) = 0$. Then $\eta^i(t_o) = 0$, $\eta^i(\tau) = -\gamma y^{i\prime}(\tau)$ and (η, μ) is not of the form

$(\rho y', \rho\lambda')$ since not all the constants A^h are zero. Consequently τ defines a point y conjugate to y_0. Suppose now that Θ changes sign without vanishing at $t = \tau$. Then $\gamma\,\Theta(\tau^-, t) + \delta\,\Theta(\tau^+, t) = 0$, where $\gamma\delta > 0$. Hence the equations

$$A^h \eta^{ih}(\tau) + \gamma_1\left[\gamma y^{i\prime}(\tau^-) + \delta y^{i\prime}(\tau^+)\right] = 0,$$
$$A^h \eta^{ih}(t_0) + \delta_1 y^{i\prime}(t_0) = 0,$$

have solutions A^h, γ_1, δ_1, not all zero. The constant γ_1 is not zero since if that were the case the determinant Θ would vanish at τ. If the constants A^h are all zero we have

$$\gamma y^{i\prime}(\tau^-) + \delta y^{i\prime}(\tau^+) = 0,$$

and the solution of the accessory equations occurring in the definition of a conjugate point can be taken arbitrarily, as we can take $\epsilon = 0$ in that case. If the constants A^h are not all zero we let

$$\eta^i = A^h\eta^{ih} + \rho y^{i\prime}, \qquad \mu^\alpha = A^h\mu^{\alpha h} + \rho\lambda^{\alpha\prime}$$

where $\rho(t_0) = \delta_1$, $\rho(\tau) = \rho(t_1) = , \ldots, = \rho(t_p) = 0$. Then (η, μ) is a solution of the accessory equations not of the form $(\rho y', \rho\lambda')$ having $\eta^i(t_0) = 0$, $\eta^i(\tau) = -\gamma_1\gamma y^{i\prime}(\tau^-) - \gamma_1\delta y^{i\prime}(\tau^+)$, where $\gamma_1^2\gamma\delta > 0$. Hence τ defines a point y that is conjugate to y_0. This completes the proof of the lemma.

Suppose we have a one-parameter family of extremaloids

$$y^i = y^i(t, b), \qquad \lambda^\alpha = \lambda^\alpha(t, b),$$

containing the extremaloids E_0 for $b = b_0$, and with all the corners occurring at the fixed parameter values $t_1, \ldots, t_p$. If the functions (y, λ) are substituted in the Euler-Lagrange equations (2:1) and the corner conditions (2:2) these become identities in t and b. If we differentiate these identities

with respect to b and set $b = b_0$ we obtain

$$\frac{d}{dt}(F_{y^{i\prime}y^j}\eta^j + F_{y^{i\prime}y^{j\prime}}\eta^{j\prime} + F_{y^{i\prime}\lambda^\alpha}\mu^\alpha) = F_{y^i y^j}\eta^j + F_{y^i y^{j\prime}}\eta^{j\prime} + F_{y^i\lambda^\alpha}\mu^\alpha,$$

$$\varphi^\alpha_{y^i}\eta^i + \varphi^\alpha_{y^{i\prime}}\eta^{i\prime} = 0,$$

$$F^-_{y^{i\prime}y^j}\eta^j + F^-_{y^{i\prime}y^{j\prime}}\eta^{j\prime -} + F^-_{y^{i\prime}\lambda^\alpha}\mu^{\alpha -} = F^+_{y^{i\prime}y^j}\eta^j + F^+_{y^{i\prime}y^{j\prime}}\eta^{j\prime +} + F^+_{y^{i\prime}\lambda^\alpha}\mu^{\alpha +},$$

where $\eta^j = y^i_b(t,b_0)$; $\eta^{j\prime} = y^{i\prime}_b(t,b_0)$; $\mu^\alpha = \lambda^\alpha_b(t,b_0)$. These equations are seen to be the same as equations (7:4) and corner conditions (7:5). Hence (η,μ) is a solution of the accessory equations. From the 2n-2 parameter family of extremaloids

$$y^i = \psi^i(t,b), \qquad \lambda^\alpha = \Lambda^\alpha(t,b)$$

discussed in section 5 we obtain in this manner a set of 2n-2 solutions of the accessory equations,

$$\eta^{ih} = \psi^i_{b^h}(t,b), \qquad \mu^{\alpha h} = \Lambda^\alpha_{b^h}(t,b),$$

which make the columns of the matrix M linearly independent. To see this we make use of the general solution

$$\text{(9:2)} \qquad \begin{aligned} y^i &= \psi^i(t,\xi,\xi',l) \\ \lambda^\alpha &= L^\alpha(t,\xi,\xi',l) \end{aligned}$$

of equations (4:3) considered in section 4.

Consider the columns of the matrix of derivatives of $y^i, y^{i\prime}, \lambda^\alpha$, with respect to $\xi^i, \xi^{i\prime}, l^\alpha$,

$$\text{(9:3)} \qquad \begin{Vmatrix} y^i_{\xi^j} & y^i_{\xi^{j\prime}} & y^i_{l^\beta} \\ y^{i\prime}_{\xi^j} & y^{i\prime}_{\xi^{j\prime}} & y^{i\prime}_{l^\beta} \\ \lambda^\alpha_{\xi^j} & \lambda^\alpha_{\xi^{j\prime}} & \lambda^\alpha_{l^\beta} \end{Vmatrix}$$

At the initial value (τ,ξ_0,ξ_0',l_0) the matrix (9:3) becomes the identity matrix as can readily be verified by differentiat-

ing equations (9:2) at the point $t = \tau$. At the initial value (ξ_0, ξ_0') the matrix (1:3) is assumed to have rank m. Hence the equations $\varphi^\alpha(\xi_0, \xi_0') = 0$ can be solved for m of the $\xi_0^{i\prime}$ in terms of the remaining n-m $\xi_0^{i\prime}$, and the ξ_0^i. Call these 2n-m initial values v. Not all the n-m $\xi_0^{i\prime}$ in v can be zero on account of the homogeneity of the functions φ^α and the consequent homogeneity of the solutions. In the matrix of derivatives (9:3) we can omit the rows and columns for the m $\xi_0^{i\prime}$ determined by the equations $\varphi^\alpha(\xi_0, \xi_0') = 0$. The remaining matrix is still the identity matrix if $t = \tau$. Now two columns of this matrix can be replaced by the last two columns of M without destroying the linear independence of the columns. That is, if $\xi_0^{r\prime} \neq 0$ we replace the columns corresponding to derivatives with respect to ξ_0^r and $\xi_0^{r\prime}$ respectively. Thus we have shown the existence of one minor of M of order 2n that does not vanish at $t = \tau$ and consequently the columns of M are linearly independent throughout (t_0, T) by lemma 9:1. As explained in the closing paragraph of section 5 the sign and the zeros of the determinant (5:3) are unaffected by the transformation necessary to make the corners appear at fixed parameter values. Lemma 9:2 combined with these facts yields the following theorem.

THEOREM. *Let $y^i = \psi^i(t,b)$, $\mu^\alpha = \Lambda^\alpha(t,b)$ be the equations of the general 2n-2 parameter family of extremaloids in which the extremaloid E_0 is contained for $b = b_0$. Then the points y conjugate to y_0 on E_0 correspond to the values of the parameter $t \neq t_0$ for which the determinant*

$$\begin{vmatrix} \dfrac{\partial\psi^i(t,b_o)}{\partial b^h} & \dfrac{\partial\psi^i(t,b_o)}{\partial t} & 0 \\ \\ \dfrac{\partial\psi^i(t_o,b_o)}{\partial b^h} & 0 & \dfrac{\partial\psi^i(t_o,b_o)}{\partial t} \end{vmatrix}$$

<u>vanishes or changes sign</u>.

A PROBLEM IN THE CALCULUS OF VARIATIONS SUGGESTED BY A PROBLEM IN ECONOMICS

BY

HENRY HOWES PIXLEY

TABLE OF CONTENTS

INTRODUCTION

The economic problem under consideration is that of determining the rate of production at which a monopolist should operate his business to furnish a maximum profit. In §1 three sets of assumptions under which he may operate will be considered and the corresponding economic problems stated. In all three cases it is assumed that the rate of production and selling price are functions of the time. In the first monopoly problem the producer fixes the price curve arbitrarily after which the demand curve is determined by a demand law of the Evans type (1·1) in which the rate of demand depends upon the price and the rate of change of price with respect to the time.[1] In the second monopoly problem the producer fixes the offer curve arbitrarily after which the price curve is determined by a price law of type (1·5) in which the price is a function of the total amount which the market has absorbed since a given time, the offer, and the rate of change of offer. The third monopoly problem is a specialization of the second, the only change being that the price is a function of offer and rate of change of offer only.

---

[1] G. C. Evans, Mathematical Introduction to Economics (New York, 1930), p. 36.

In all three cases the cost function (1·2) is assumed to be a function of the rate of production and the time.

The first and third problems are special cases of the problem in the calculus of variations of maximizing an integral of type $\int_{x_1}^{x_2} f(x,y,y')dx$ in a class of arcs $y=y(x)$ such that $y(x)$ and $y'(x)$ may have an arbitrary finite number of finite discontinuities. The maximizing arc is to join two given vertical lines in the xy-plane. The possible discontinuities in $y(x)$ come from the assumption that a discontinuity may be introduced into the price function or rate of production function at any time.

This calculus of variations problem is considered in § 2. It is found that there exists no solution unless the two differential equations $f_y(x,y,y')=f_{y'}(x,y,y')=0$ have a solution in common. Analogues of the Weierstrass and Legendre necessary conditions and a set of conditions which insure a weak relative maximum are derived. The results of this problem are applied in § 3 to a special case of the first monopoly problem. This problem differs from one treated by Evans only in the introduction of the assumption which allows discontinuities in the price and production curves.[1] It is found for this special case that no price curve maximizing the profit exists unless the demand function is indepenaent of the

---

[1] Ibid., p. 143.

rate of change of price. If this is true a maximizing arc may exist. If so, the maximizing function is continuous and is a constant. Furthermore, it makes the integral an absolute maximum. These results are also established for a slightly more general type of demand function. Analogous results are found for the third monopoly problem.

The second monopoly problem is a special case of the problem of maximizing an integral of the type $\int_{x_1}^{x_2} f(x,y,y',y'')dx$ in a class of continuous arcs $y=y(x)$ such that $y'(x)$ and $y''(x)$ may have an arbitrary finite number of finite discontinuities. The arcs of this class are also required to join a fixed point to a given vertical line. This calculus of variations problem is generalized slightly by replacing the given vertical line by a general curve of class C''. It is treated in §§4-7. Again it is found that a maximizing arc must satisfy two differential equations which ordinarily have no solution in common. This condition and a transversality condition are derived in § 4. Analogues of the Weierstrass and Legendre necessary conditions are obtained in § 5. Several analogues of the Jacobi condition are obtained and compared in §6. Two of these conditions are expressible in terms of envelope theorems based on a one-parameter family of "extremals" in the xy-plane. In §7 the necessary conditions of §§4-6 are strengthened to give a set of conditions which insure a strong relative maximum. The conclusions which can be obtained by apply-

ing the mathematical theory of §§4-6 to the second monopoly problem are for the most part quite analogous to those obtained in §3 when the third monopoly problem was studied. A formal example of the problem of §4 for which a maximum does exist is given in §8. It is not possible to use the second monopoly problem for an illustration since it does not satisfy the sufficient conditions of §7.

The results obtained for the first monopoly problem seem to indicate that one or more of the following possibilities exist; the rate of change of price with respect to the time should not enter into the demand function, the forms of the cost and demand functions used in the present formulation of the problem are not suitable, or the comparison price curves should not be allowed to have actual discontinuities. Of these possible objections to the present form of the problem it seems easier to sustain the second than either of the others. This suggests that further investigation of this problem should either search for more satisfactory cost and demand functions or else reformulate the problem so that it falls under a different mathematical theory. Similar remarks apply also to the second and third monopoly problems.

1. The monopoly problems. Throughout the discussion of the monopoly problems the independent variable t will represent the time. The three functions $p(t)$, $u(t)$, $y(t)$ will represent the price per unit, the rate of production, and the

rate of demand, respectively. It is assumed that the interval of time from t_1 to t_2 can be divided into a finite number of intervals such that $p(t)$ is of class C' in each of these intervals, and both the right and left limits of $p(t)$ and $p'(t)$ exist for every value of t in the closed interval (t_1, t_2). The functions $u(t)$ and $y(t)$ possess the same properties.

In the first monopoly problem the producer fixes the price $p(t)$ and produces the amount which he can sell at this price, that is, $u(t) = y(t)$. It is assumed that the rate of demand $y(t)$ depends upon the price, the rate of change of price p', and the time. This is expressed by the demand equation

$$y = y(p, p', t). \tag{1.1}$$

The cost per unit time $Q(t)$ of maintaining a rate of production $u(t)$ is given as a function of u and t.

$$Q(t) = Q[u(t), t] \tag{1.2}$$

It is true in general that at a time t at which the price function is discontinuous the subsequent values of the rate of demand will depend not only upon the new value of the price but also upon the change of price, that is, y will be a function of the discontinuity. A similar phenomenon is true for the function Q when u has a discontinuity. Also the rate of demand may depend upon the history of the price in other ways than the one just mentioned. Hence the following theory is valid only for those problems in which these effects are negligible when com-

pared with those which have been introduced in equations (1·1) and (1·2).

Under these assumptions the gross income and the profit per unit time at the time t are $p(t)y(t) = p(t)u(t)$ and $p(t)u(t) - Q(t)$, respectively, and the total value of the profits in the time interval from t_1 to t_2 is given by the integral[1]

$$(1\cdot3) \qquad \Pi = \int_{t_1}^{t_2}(pu-Q)dt.$$

The producer may not wish to maximize this total value of his profit. He may wish to maximize the total value of the profit for the interval (t_1, t_2) after it has been discounted to a time T. The value $p(t)u(t) - Q(t)$ of the profit per unit time at time t when discounted to time T is $(pu-Q)E(T,t)$ where $E(T,t) = \exp\left[-\int_T^t \delta(t)dt\right]$ is a discount (or interest) factor. The function $\delta(t)$ is the force of interest, or rate of increase of an invested sum S divided by S.[2] Under this assumption it is the integral

$$(1\cdot4) \qquad \Pi = \int_{t_1}^{t_2}(pu-Q)E(T,t)dt$$

which the producer wishes to maximize. The time T may be a time at which the producer wishes to reorganize his business. If t_1 represents "now" and he wishes to maximize the present

---

[1]See Evans, loc. cit., p. 144.

[2]See H. Hotelling, A general mathematical theory of depreciation, Journal of the American Statistical Association (September, 1925), p. 342.

value of his profit then $T = t_1$. If $\delta(t) = 0$, the integral (1·4) becomes the integral (1·3). Hence (1·3) is a special case of (1·4).

In the second monopoly problem the producer fixes his rate of production $u(t)$ and offers his entire production on the open market. It is assumed that the price depends upon the offer u, the rate of change of offer u', the time t, and the total amount that has been offered since the time t_1, that is, $\int_{t_1}^{t} u\,dt$. This is expressed by the <u>price equation</u>[1]

$$p = p\left(\int_{t_1}^{t} u\,dt, u, u', t\right). \tag{1·5}$$

The cost function (1·2) which was used before is used here also with the same assumptions involved. It is further assumed that the changes in price caused by the discontinuities of u are negligible when compared with those effects which have been introduced in the function (1·5). As before the producer wishes to maximize the integral (1·4).

The third monopoly problem is a specialization of the second. Suppose all the assumptions made in the second problem hold except that the price equation takes the form

$$p = p(u, u', t), \tag{1·6}$$

that is, the price does not depend upon the amount that has been offered since time t_1.

---

[1]For a price function of this type in a monopoly problem of another kind see H. Hotelling, The economics of exhaustible resources, The Journal of Political Economy, XXXIX (1931), pp. 152, 153.

The first and third problems require the same mathematical analysis. This analysis will be made in §2 and applied to these cases in §3. The second monopoly problem requires a different analysis which will be made in §§4-8.

SOLUTIONS WHICH ARE ACTUALLY DISCONTINUOUS

IN THE PROBLEM WITH FIRST DERIVATIVES

2. Necessary conditions and sufficient conditions for a maximum. Statement of the problem. Consider a class of admissible arcs of the form

$$y = y(x), \qquad (x_1 \leqq x \leqq x_2), \tag{2.1}$$

each of which joins the line $x = x_1$ to the line $x = x_2$. We wish to find in this class of arcs one which maximizes the integral

$$I = \int_{x_1}^{x_2} f(x,y,y')dx.$$

It is assumed that there is a region R of sets of values (x,y,y') in which the function $f(x,y,y')$ and all of its partial derivatives of the first two orders are continuous. A set (x,y,y') which is interior to R will be called an admissible set. An admissible function $y(x)$ is one which possesses the properties, (a) $y(x)$ and its derivative $y'(x)$ are each continuous except possibly for a finite number of values of x, (b) for every value of x in the interval of definition the right and left hand limits of $y(x)$ and $y'(x)$ exist, (c) the sets $[x,y(x),y'(x)]$ belonging to the function are all admissible. Wherever it is convenient the right and left hand

limits of $y(x)$ and $y'(x)$ will be denoted by $y_+(x)$, $y'_+(x)$ and $y_-(x)$, $y'_-(x)$, respectively. The element $[x, y(x), y'(x)]$ shall be interpreted to mean either $[x, y_-(x), y'_-(x)]$ or $[x, y_+(x), y'_+(x)]$. An arc $y = y(x)$ defined by an admissible function will be called an <u>admissible arc</u>. It will be noticed that the integral I has a well-defined value over every admissible arc. A function $\eta(x)$ which has properties (a) and (b) will be called an <u>admissible variation</u>.

<u>Analogues of the Weierstrass, Euler, and Legendre necessary conditions</u>. Suppose the equation

$$E_{34}: \qquad y = y(x), \qquad (x_3 \leqq x \leqq x_4),$$

defines a portion of a maximizing arc, and the equation

$$C_{3'4'}: \qquad y = Y(x), \qquad (x_3 \leqq x \leqq x_4),$$

defines any admissible arc in the interval (x_3, x_4). Let x_5 be any value between x_3 and x_4 and designate the points which it defines on the arcs E_{34} and $C_{3'4'}$ by 5 and 5', respectively. Since the path $C_{3'5'} + E_{54}$ constitutes an admissible arc, the integral

$$I(C_{3'5'} + E_{54}) = \int_{x_3}^{x_5} f(x, Y, Y')dx + \int_{x_5}^{x_4} f(x, y, y')dx$$

must be less than or equal to $I(E_{34})$ for every value of x_5 between x_3 and x_4. Hence the derivative dI/dx_5 must not be positive when $x_5 = x_3$. This gives the condition

$$\frac{dI}{dx_5}\bigg|^{x_5 = x_3} = f[x_3, Y(x_3), Y'(x_3)] - f[x_3, y(x_3), y'(x_3)] \leqq 0.$$

Since the arc $C_{3'4'}$ may be any admissible arc this inequality must hold for an arbitrary admissible set $[x_3, Y(x_3), Y'(x_3)]$.

ANALOGUE OF THE WEIERSTRASS NECESSARY CONDITION. <u>At every element</u> (x,y,y') <u>of a maximizing arc</u> E_{12} <u>the condition</u>

$$f(x,Y,Y') - f(x,y,y') \leqq 0 \tag{2·2}$$

<u>must hold for every admissible set</u> $(x,Y,Y') \neq (x,y,y')$.

This condition implies that for every value of x the pair of values (y,y') associated with the maximizing arc furnish a maximum for $f(x,y,y')$ as compared with any other values (Y,Y') such that (x,Y,Y') is an admissible set. This implies, as a result of the continuity properties of f, that at every element (x,y,y') of a maximizing arc E_{12} the conditions

$$f_y(x,y,y') = f_{y'}(x,y,y') = 0 \tag{2·3}$$

$$f_{y'y'}(x,y,y') \leqq 0, \quad f^2_{yy'}(x,y,y') - f_{yy}(x,y,y')f_{y'y'}(x,y,y') \leqq 0, \tag{2·4}$$

$$f_{yy}(x,y,y') \leqq 0,$$

must be satisfied. The pair of equations (2·3) is the analogue of the Euler equation and the system of inequalities (2·4) is the analogue of the Legendre condition. Since ordinarily there will be no solution common to the two equations (2·3) it is true that in general among the admissible arcs (2·1) which join the line $x = x_1$ to the line $x = x_2$ none gives the integral I a maximum value. If such a maximizing arc E_{12} does exist it is defined by equations (2·3) and satisfies conditions (2·2) and (2·4). Equations (2·3) are sufficient to make the first variation of the integral I, defined in the

usual way,[1] vanish. Hence no further necessary conditions can be found by requiring the first variation to vanish.

At a place on a maximizing arc E_{12} where either y or y' is discontinuous there are two elements (x, y_-, y'_-) and (x, y_+, y'_+) each associated with the same value of x. It follows from the Weierstrass condition that $f(x, y_-, y'_-) = f(x, y_+, y'_+)$.

ANALOGUE OF THE CORNER CONDITION. <u>At every element of a maximizing arc the equation</u>

$$f(x, y_-, y'_-) = f(x, y_+, y'_+)$$

<u>must be satisfied.</u>

<u>Sufficient conditions</u>. Let E_{12} be an arc of class C' which satisfies equations (2·3). For an element (x, y, y') on E_{12} at which the inequalities (2·4) hold in the strengthened form, that is, without the equality sign, the inequality (2·2) also holds in the strengthened form for values (Y, Y') near (y, y'). Hence, if the inequalities (2·4) hold in the strengthened form for every element (x, y, y') of E_{12} there exists a region R' of sets (x, Y, Y') near the elements (x, y, y') of E such that the inequality $f(x, Y, Y') - f(x, y, y') < 0$ holds for every set (x, Y, Y') in R' which is not an element of E_{12}. Then if the equation $y = Y(x), x_1 \leqq x \leqq x_2$, represents an admissible arc all of whose elements $[x, Y(x), Y'(x)]$ are in R', but which

---

[1]See Bolza, Vorlesungen über Variationsrechnung (Leipzig, 1909), p. 21.

is not identical with E_{12}, it follows that

$$I(C_{12}) - I(E_{12}) = \int_{x_1}^{x_2} \left[f(x,Y,Y') - f(x,y,y')\right] dx < 0.$$

If y or y' have finite discontinuities this inequality holds if every set $[Y(x), Y'(x)]$ is sufficiently close to the set $[y(x), y'(x)]$ of E_{12} defined by the same value of x.

SUFFICIENT CONDITIONS FOR A WEAK RELATIVE MAXIMUM.

Let E_{12} be an admissible arc which satisfies the equations $f_y(x,y,y') = f_{y'}(x,y,y') = 0$ and has $f_{y'y'} < 0$, $f_{yy'}^2 - f_{yy} f_{y'y'} < 0$ at every element (x,y,y') on it; then the inequality $I(C_{12}) < I(E_{12})$ holds for every admissible arc C_{12} which does not coincide with E_{12}, and which is such that every set $[Y(x), Y'(x)]$ belonging to it lies sufficiently near the set $[y(x), y'(x)]$ of E_{12} defined by the same value of x.

3. An application to the first and third monopoly problems. A special case of the first monopoly problem is characterized by demand and cost equations, (1·1) and (1·2), which take the forms

$$y = ap + bp' + c, \tag{3·1}$$

$$Q = Au^2 + Bu + C, \tag{3·2}$$

respectively, where a,b,c,A,B,C are constants.[1] In this case the integral (1·4) is

---

[1]See Evans, loc. cit., p. 143.
See C. F. Roos, A mathematical theory of depreciation and replacement, American Journal of Mathematics, Vol. L(1928) for an extension of this problem.

$$\Pi = \int_{t_1}^{t_2} f(t,p,p')dt \tag{3·3}$$

where $f(t,p,p') = \left[a_{00}p^2 + a_{01}pp' + a_{11}p'^2 + a_0 p + a_1 p' + d\right]E(T,t)$,

$$a_{00} = a(1-aA), \qquad a_{01} = b(1-2aA). \qquad a_{11} = -b^2 A,$$

$$a_0 = c-2acA - aB, \; a_1 = -b(B+2cA), \quad d = -(Ac^2 + Bc + C).$$

The function $f(t,p,p')$ is continuous with continuous partial derivatives of all orders for every set (t,p,p'). Hence the maximizing of the integral (3·3) in the class of arcs,

$$p = p(t), \qquad (t_1 \leqq t \leqq t_2),$$

which are admissible in the sense of §2 and join the line $t = t_1$ to the line $t = t_2$, becomes a special case of the problem of §2. The equations (2·3) of the maximizing arc become

$$f_{p'} = \left[2a_{11}p' + a_{01}p + a_1\right] E(T,t) = 0 \tag{3·4}$$

$$f_p = \left[a_{01}p' + 2a_{00}p + a_0\right] E(T,t) = 0. \tag{3·5}$$

Since the factor $E(T,t)$ is greater than zero for every value t it may be divided from these equations. It follows from the analogue (2·4) of the Legendre condition that along an arc which furnishes a maximum the inequalities $f_{p'p'} = 2a_{11}E(T,t) \leqq 0$, $f_{pp'}^2 - f_{pp}f_{p'p'} = (a_{01}^2 - 4a_{00}a_{11})E(T,t) \leqq 0$ hold. These inequalities imply $a_{11} \leqq 0$, $a_{01}^2 - 4a_{00}a_{11} \leqq 0$. In terms of the coefficients of the functions (3·1) and (3·2) the second of these inequalities is $a_{01}^2 - 4a_{00}a_{11} = b^2 \leqq 0$. Therefore, there exists no arc which will maximize the integral (3·3) unless $b = 0$. This implies $a_{01} = a_{11} = a_1 = 0$ and equations (3·4) and (3·5) reduce to $2a_{00}p + a_0 = 0$. If $a_{00} < 0$, the equation of the maximizing arc is

$$(3\cdot6) \qquad p = -\frac{a_o}{2a_{oo}}.$$

If f satisfies the above conditions it may be written $f(t,p) = [a_{oo}p^2 + a_o p + d]E(T,t)$. It is evident that the value (3·6) gives this function an absolute maximum in p, that is, $f(t,P) - f(t,p) < 0$ for every t, and every $P \neq p = -a_o/2a_{oo}$. If the arc (3·6) is designated by E_{12} and the equation $p = P(t)$ represents any other admissible arc C_{12} which joins $t = t_1$ to $t = t_2$, it follows that

$$I(C_{12}) - I(E_{12}) = \int_{t_1}^{t_2} \{f[t,P(t)] - f[t,p]\}\,dt < 0.$$

Hence the arc (3·6) furnishes an absolute maximum in the class of admissible arcs joining $t = t_1$ to $t = t_2$.

The foregoing results hold for demand equations more general than (3·1) when the cost function (3·2) is used. Suppose

$$(3\cdot7) \qquad y = ap + bP(p') + c$$

where $P(p')$ is of class C' for all values of p', $P'(p')$ has no zeros, and $P(0) = 0$.[1] In this case the inequalities (2·4) imply $f_{pp'}^2 - f_{pp}f_{p'p'} = (a_{o1}^2 - 4a_{oo}a_{11})P'^2(p')E(T,t) = b^2P'^2(p')E(T,t) \leqq 0$. Hence $b = 0$ and the proof used before proves that the arc (3·6) furnishes an absolute maximum.

Therefore, for a business operated under the demand and cost functions (3·1) and (3·2) with $b \neq 0$, there exists

---

[1]See Evans, loc. cit., p. 42.

no price function $p(t)$ which will maximize the profit integral π in the class of all possible price functions if these comparison price functions are free to have actual discontinuities at any finite number of places. This, of course, does not mean that the mathematical analysis forbids the use of the derivative p' in the demand equation in this problem. It merely means that if p' is to be introduced into the demand equation the form (3·1) is not suitable. Furthermore the generalization (3·7) is unsuitable. There are forms of the functions (1·1) and (1·2) which lead to an integral (1·4) which has p' in the integrand and satisfies the sufficient conditions of § 2. To determine which of these forms has any significance in the economic problem is a major problem in itself.

In the third monopoly problem the integrand function of the profit integral π is

$$f(t,u,u') = \left[up(u,u',t) - Q(u,t)\right]E(T,t).$$

The maximizing arc $u = u(t)$ must satisfy the equations

$$(3\cdot 8)\qquad \begin{aligned} f_{u'} &= up_{u'}(u,u',t)E(T,t) = 0 \\ f_u &= \left[up_u + p - Q_u\right]E(T,t) = 0. \end{aligned}$$

The first of these equations implies $u \equiv 0$ or $p_{u'}(u,u',t) = 0$. Consider the case in which $u \equiv 0$ is not a solution of the problem. Then, if the problem has a solution, $p_{u'}(u,u',t) = 0$. If the price equation takes the form

$$(3\cdot 9)\qquad p = au + bU(u') + c$$

where $U(u')$ is of class C' for all values of u', $U'(u')$ has no

zeros, and a,b,c are constants, we have

$$p_{u'} = bU'(u') = 0.$$

This equation is impossible unless b = 0. Hence the price function (3·9) is unsatisfactory in the third monopoly problem in the same sense that the demand function (3·7) is unsatisfactory in the first problem. The price function (3·9) is of interest since a special case of it,

$$p = au + bu' + c,$$

is analogous to the demand function (3·1).

If u is not identically zero, then at every element (t,u,u') on the maximizing arc at which the condition

$$f_{u'u'} = p_{u'u'}\,uE(T,t) < 0$$

holds, this condition, together with the equation $p_{u'} = 0$, implies that the value u' maximizes the function p among the sets (t,u,U') where U' is arbitrary near u'. This follows from the condition that along an admissible arc for this problem $u \geqq 0$. This conclusion holds for all forms of the price function (1·6).

SOLUTIONS WITH CORNERS IN THE PROBLEM WITH SECOND DERIVATIVES

4. First necessary conditions. If in the second monopoly problem of §1 the transformation $x = t$, $y = \int_{t_1}^{t} u\,dt$ is made the problem becomes a special case of the problem of finding in a class of admissible arcs of the form

$$y = y(x) \tag{4.1}$$

which join a given point 1, (x_1, y_1), to a given curve N, one which maximizes the integral

$$I = \int_{x_1}^{x_2} f(x,y,y',y'')dx,$$

where x_2 is the abscissa of the point of intersection of the curve $y = y(x)$ and N. It is assumed that there is a region R of sets of values (x,y,y',y'') in which the function $f(x,y,y',y'')$ and all of its partial derivatives of the first four orders are continuous. A set (x,y,y',y'') interior to R will be called an admissible set. A function $y(x)$ will be called admissible if (a) $y(x)$ is continuous, (b) $y'(x)$ and $y''(x)$ are continuous except possibly for a finite number of values of x, (c) for every value of x in the interval of definition the right and left hand limits of $y'(x)$ and $y''(x)$ exist, (d) the sets $[x,y(x),y'(x),y''(x)]$ belonging to the function are all admissible. Right and left hand limits of

the derivatives $y'(x), y''(x)$ will be denoted by $y'_+(x), y''_+(x)$ and $y'_-(x), y''_-(x)$, respectively. The element $[x, y(x), y'(x), y''(x)]$ shall be interpreted to mean either $[x, y(x), y'_-(x), y''_-(x)]$ or $[x, y(x), y'_+(x), y''_+(x)]$. An <u>admissible arc</u> is an arc $y = y(x)$ defined by an admissible function. An <u>admissible variation</u> $\eta(x)$ is a function which has properties (a), (b), (c).

Suppose

$$E_{12}: \qquad y = y(x), \qquad (x_1 \leqq x \leqq x_2),$$

is a maximizing arc for the problem under consideration. Let $\eta(x)$ be an admissible variation. Then if the parameter <u>a</u> is small the arc

$$y = y(x,a) = y(x) + a\eta(x) \tag{4·2}$$

is admissible. Let the curve N be represented by the equations

$$N: \qquad x = \phi(\alpha), \qquad y = \psi(\alpha)$$

where $\phi(0) = x_2$, $\psi(0) = y(x_2) = y_2$, $\phi'^2(0) + \psi'^2(0) \neq 0$, and ϕ, ψ are of class C'' near $\alpha = 0$. If the arc E_{12} is not tangent to N at the point 2 then the value of α which defines the point of intersection of the arc (4·2) and N can be obtained by solving the equation

$$y[\phi(\alpha)] + a\eta[\phi(\alpha)] = \psi(\alpha).$$

The function $\alpha(a)$ defined by this equation is of class C'' near $a = 0$.[1] Consequently the equation may be differentiated twice

[1] G. A. Bliss, Princeton Colloquium Lectures (New York, 1913), pp. 8, 11.

with respect to a, giving the equations

$$(4\cdot3)\qquad y'(x_2)\phi'(0)\alpha'(0) + \eta(x_2) = \psi'(0)\alpha'(0),$$

$$(4\cdot4)\qquad y''(x_2)\phi'^2(0)\alpha'^2(0) + y'(x_2)\phi''(0)\alpha'^2(0) + y'(x_2)\phi'(0)\alpha''(0) + 2\eta'(x_2)\phi'(0)\alpha'(0) = \psi''(0)\alpha'^2(0) + \psi'(0)\alpha''(0)$$

for $a = 0$.

If $\eta(x_1) = 0$ then as a varies near zero equation (4·2) represents a one-parameter family of admissible comparison arcs and the integral I taken over one of these arcs has a value which depends upon a:

$$(4\cdot5)\qquad I(a) = \int_{x_1}^{\phi[\alpha(a)]} f\left[x, y(x,a), y'(x,a), y''(x,a)\right] dx.$$

Since by hypothesis I(0) is a maximum it follows that $I'(0)=0$ for any admissible variation $\eta(x)$ having $\eta(x_1) = 0$. The <u>first variation</u> $I'(0)$ has the form

$$I'(0) = \int_{x_1}^{x_2}\left[f_y\eta + f_{y'}\eta' + f_{y''}\eta''\right]dx + f\left[x_2, y_2, y'(x_2), y''(x_2)\right]\phi'(0)\alpha'(0)$$

where the arguments of f_y, $f_{y'}$, $f_{y''}$ are the elements (x, y, y', y'') of E_{12}. Define $M(x)$ by the equation

$$M(x) = \int_{x_1}^{x}\int_{x_1}^{x} f_y\left[x, y(x), y'(x), y''(x)\right]dx^2 - \int_{x_1}^{x} f_{y'}\left[x, y(x), y'(x), y''(x)\right]dx + f_{y''}\left[x, y(x), y'(x), y''(x)\right].$$

If $\eta'(x)$ is continuous except possibly for $x = x_3$ the first variation can be transformed by an integration by parts into

the form

$$I'(0) = \eta(x_2)\int_{x_1}^{x_2} f_y\,dx + \eta'(x_2)\left[\int_{x_1}^{x_2} f_{y'}\,dx - \int_{x_1}^{x_2}\int_{x_1}^{x} f_y\,dx^2\right]$$

$$(4\cdot 6) \qquad - \left[\eta'_+(x_3) - \eta'_-(x_3)\right]\cdot\left[\int_{x_1}^{x_3} f_{y'}\,dx - \int_{x_1}^{x_3}\int_{x_1}^{x} f_y\,dx^2\right]$$

$$+ \int_{x_1}^{x_2} M(x)\eta''\,dx + f\left[x_2, y_2, y'(x_2), y''(x_2)\right]\phi'(0)\,\alpha'(0)\,.$$

In studying this expression the admissible variation $\eta(x)$ may have any further restrictions placed upon it that we find helpful. If for instance $\eta(x_1) = \eta'(x_1) = \eta(x_2) = \eta'(x_2) = 0$ and η' is continuous for all values of x then

$$(4\cdot 7) \qquad I'(0) = \int_{x_1}^{x_2} M(x)\eta''\,dx$$

The term $f\left[x_2, y_2, y'(x_2), y''(x_2)\right]\phi'(0)\alpha'(0)$ drops out of the first variation in this case as a result of equation (4·3) and the assumption that E_{12} and N are not tangent at the point 2. A necessary condition for our problem is that the expression (4·7) should vanish for every choice of the admissible variation $\eta(x)$ which is of class C', and such that $\eta(x_1) = \eta'(x_1) = \eta(x_2) = \eta'(x_2) = 0$. It is known that if the expression (4·7) vanishes for all η's of this kind which satisfy the additional condition that η'' is continuous then the function $M(x)$ is necessarily of the form

$$(4\cdot 8) \qquad M(x) = c_1(x - x_1) + c_2$$

where c_1 and c_2 are constants.[1] Furthermore if this expression is substituted for $M(x)$ in the integral (4·7) it is clear that that integral vanishes even if the condition that η'' be continuous is removed.

Suppose next that η vanishes together with its first derivative for $x = x_1$ and $x = x_2$ and that η' is continuous except for a value $x = x_3$ between x_1 and x_2. In this case, if equation (4·8) is satisfied, the first variation (4·6) takes the form

$$I'(0) = -\left[\eta'_+(x_3)-\eta'_-(x_3)\right]\cdot\left[f_{y''}(x_3)-c_1(x_3-x_1)-c_2\right]$$
$$+\int_{x_1}^{x_2}\eta''\left[c_1(x-x_1)+c_2\right]dx = -\left[\eta'_+(x_3)-\eta'_-(x_3)\right]f_{y''}(x_3).$$

Here $f_{y''}(x_3) = f_{y''}\left[x_3, y(x_3), y'(x_3), y''(x_3)\right]$ where either of the sets $\left[y'_-(x_3), y''_-(x_3)\right]$, $\left[y'_+(x_3), y''_+(x_3)\right]$ can be used for $\left[y'(x_3), y''(x_3)\right]$ since it follows from equation (4·8) that $f_{y''}$ is a continuous function of x along the arc E_{12}. The vanishing of $I'(0)$ in this last form implies that $f_{y''}(x_3) = 0$. Since for any value of x_3 between x_1 and x_2 there will be an admissible variation $\eta(x)$ in the class of η's under consideration which will have a discontinuous derivative at x_3, it follows that $f_{y''} = 0$ at every element (x, y, y', y'') of E_{12}.[2] Our results so far may be stated as the

---

[1]See E. Zermelo, Mathematische Annalen, Vol. LVIII (1904), p. 558.

[2]See Hadamard, Lecons sur le Calcul des Variations (Paris, 1910), pp. 145, 146.

LEMMA. For every arc E_{12} which furnishes a maximum for the integral I there exists a constant c such that the equations

$$(4\cdot 9) \qquad \int_{x_1}^{x}\int_{x_1}^{x} f_y\,dx\cdot dx - \int_{x_1}^{x} f_{y'}dx = c(x-x_1)$$

$$(4\cdot 10) \qquad f_{y''} = 0$$

are satisfied identically along E_{12}.

An immediate corollary of this lemma is that the equations

$$(4\cdot 11) \qquad \begin{aligned} \int_{x_1}^{x} f_y\,dx - f_{y'} &= c \\ f_y - \frac{d}{dx} f_{y'} &= 0 \end{aligned}$$

must be satisfied identically along every portion of E_{12} which is of class C'''. If we define corner to mean a point on a curve where either the first or second derivative of the representing function is discontinuous, we can obtain the following condition from the first of equations (4·11).

FIRST CORNER CONDITION. At a corner (x,y) of a maximizing arc the condition

$$f_{y'}(x,y,y'_-,y''_-) = f_{y'}(x,y,y'_+,\ y''_+)$$

must hold.

Ordinarily there will be no arc which satisfies both of equations (4·9) and (4·10). Hence ordinarily there will be no maximizing arc for the integral I in the class of arcs under consideration. However an example will be exhibited

for which a maximizing arc does exist among the admissible arcs.

If E_{12} satisfies equation (4·9) and (4·10) the first variation (4·6) takes the form

$$I'(0) = \left\{\left[\int_{x_1}^{x_2} f_y dx - c\right]\cdot\left[\psi'(0) - y'(x_2)\phi'(0)\right] + f(x_2)\phi'(0)\right\}a'(0).$$

In this equation $\eta(x_2)$ has been replaced by its value obtained from equation (4·3) and $f(x_2) = f\left[x_2, y_2, y'(x_2), y''(x_2)\right]$. As a consequence of equation (4·3) the quantity $a'(0)$ cannot vanish if $\eta(x_2) \neq 0$. Hence the first factor of $I'(0)$ in this last form must vanish. This gives, with the use of the first of equations (4·11), the equation

$$(4\cdot 12) \quad \left[f(x_2) - y'(x_2)f_{y'}(x_2)\right]\phi'(0) + f_{y'}(x_2)\psi'(0) = 0.$$

If this equation is satisfied the arc N is said to cut the arc E_{12} transversally at the point 2. The conditions which have been obtained by requiring the first variation to vanish may be summarized as follows.

THEOREM. In general none of the admissible arcs (4·1) gives the integral I a value at least as great as that given by any other admissible arc. If such a maximizing arc E_{12} exists then the equations (4·9) and (4·10) are satisfied identically along it. Furthermore E_{12} must satisfy the transversality condition (4·12).

By choosing a as independent parameter in place of a in the study of the integral (4·5) the transversality condition (4·12) can be obtained in the case where E_{12} and N are tangent at the point 2.

5. <u>The analogues of the Weierstrass and Legendre necessary conditions.</u> An arc of class C''' which satisfies the two differential equations

$$(5\cdot1) \qquad f_y(x,y,y',y'') - \frac{d}{dx} f_{y'}(x,y,y',y'') = 0$$

$$(5\cdot2) \qquad f_{y''}(x,y,y',y'') = 0$$

will be called an <u>extremal</u>. An arc of class C''' which satisfies equation (5·2) will be called a <u>semi-extremal</u>. By the theorem of § 4 every portion of a maximizing arc which is of class C''' must be an extremal.

Consider a one-parameter family of admissible arcs

$$E_b: \qquad y = y(x,b), \qquad [x_3(b) \leqq x \leqq x_4(b)],$$

which includes for the parameter value $b = 0$ a portion

$$E_{34}: \quad y = y(x,0) = y(x), \quad [x_3(0) = x_3 \leqq x \leqq x_4 = x_4(0)],$$

of an extremal arc E_{12}. Let the functions $x_3(b)$, $x_4(b)$, $y(x,b)$, $y'(x,b)$, $y''(x,b)$ be continuous with continuous first partial derivatives near $b = 0$. They define the curves

$$C: \qquad x = x_3(b), \qquad y = y[x_3(b),b] = y_3(b),$$

$$D: \qquad x = x_4(b), \qquad y = y[x_4(b),b] = y_4(b).$$

The differentials along these curves and the slope of the corresponding curve E_b at the same point are related by the equations

$$(5\cdot3) \qquad \begin{aligned} dy_3 &= y'[x_3(b),b]\,dx_3 + y_b[x_3(b),b]\,db, \\ dy_4 &= y'[x_4(b),b]\,dx_4 + y_b[x_4(b),b]\,db. \end{aligned}$$

The differential of the integral

$$I(E_b) = \int_{x_3(b)}^{x_4(b)} f[x,y(x,b),y'(x,b),y''(x,b)]\,dx$$

taken over the arc E_b is

$$dI(E_b) = f(x,b)dx\Big|_{x_3(b)}^{x_4(b)} + \left\{\int_{x_3(b)}^{x_4(b)} \left[f_y(x,b)y_b + f_{y'}(x,b)y_b' + f_{y''}(x,b)y_b''\right]dx\right\}db$$

where $f(x,b) = f\left[x, y(x,b),\ y'(x,b),\ y''(x,b)\right]$ and $f_y(x,b)$, $f_{y'}(x,b)$, $f_{y''}(x,b)$ are defined similarly. By hypothesis E_{34} is an extremal. Hence

$$f_y(x,0) - \frac{d}{dx} f_{y'}(x,0) = 0,$$
$$f_{y''}(x,0) = 0,$$

and the differential of the integral taken at E_{34} is

$$dI(E_{34}) = f(x,y,y',y'')dx\Big|_{x_3}^{x_4} + \left\{\int_{x_3}^{x_4} \left[y_b \frac{d}{dx} f_{y'} + y_b' f_{y'}\right]dx\right\}db$$

$$= \Big[f(x,y,y',y'')dx + f_{y'}(x,y,y',y'')y_b\,db\Big]_{x_3}^{x_4}.$$

By the use of equations (5·3) this becomes

$$(5\cdot4)\quad dI(E_{34}) = \Big[f(x,y,y',y'')dx + f_{y'}(x,y,y',y'')(dy - y'dx)\Big]_3^4$$

where the differentials dx, dy belong to the curves C and D and the elements (x,y,y',y'') belong to E_{34} at the points defined by x_3 and x_4, respectively.

Let 3 and 4 be points on a maximizing arc E_{12} defined by x_3 and x_4 such that $x_3 < x_4$ and let there be no corners on E_{34}. Suppose the curve

C: $\qquad y = Y(x)$

is an arbitrary curve of class C^{*} which passes through the point 3. Join the points $[x_5,\ Y(x_5)]$ of this curve to the point 4 by a family of admissible arcs

$$y = y(x,x_5), \qquad (x_5 \leqq x \leqq x_4),$$

containing E_{34} for $x_5 = x_3$ and having $y(x,x_5), y'(x,x_5)$, $y''(x,x_5)$ continuous with continuous partial derivatives of the first order. One such family is

$$y(x,x_5) = y(x) + \frac{Y(x_5) - y(x_5)}{x_4 - x_5}(x_4 - x).$$

This family, which includes E_{34} for the parameter value $x_5 = x_3$, is of the type used in deriving the formula (5·4). In this case the curve D degenerates into the point 4. Hence $dx_4 = dy_4 = 0$ and an application of the formula (5·4) gives

$$(5\cdot5)\qquad dI(E_{34}) = -\Big[f(x,y,y',y'')dx + f_{y'}(x,y,y',y'')(dY - y'dx)\Big]^{3}.$$

Consider the value of I taken along the path $C_{35} + E_{54}$. This must not be greater than $I(E_{34})$ for any value of x_5. Hence the differential of

$$I(C_{35} + E_{54}) = \int_{x_3}^{x_5} f(x,Y,Y',Y'')dx + \int_{x_5}^{x_4} f\big[x,y(x,x_5),y'(x,x_5),y''(x,x_5)\big]\,dx$$

must not be positive when $x_5 = x_3$. Using formula (5·5) we find this differential to be

$$f\big[x_3,y_3,Y'(x_3),Y''(x_3)\big]dx - f\big[x_3,y_3,y'(x_3),y''(x_3)\big]dx$$
$$- f_{y'}\big[x_3,y_3,y'(x_3),y''(x_3)\big]\cdot\big[dY - y'dx\big]$$

where $y_3 = y(x_3)$.

ANALOGUE OF THE WEIERSTRASS NECESSARY CONDITION. At every element (x,y,y',y'') of a maximizing arc E_{12} the condition
$E(x,y,y',y'',Y',Y'')=f(x,y,Y',Y'')-f(x,y,y',y'')-(Y'-y')f_{y'}(x,y,y',y'') \leqq 0$
must hold for every admissible set $(x,y,Y',Y'') \neq (x,y,y',y'')$.
In particular, if $Y'=y'$ the condition $f(x,y,y',Y'')-f(x,y,y',y'')\leqq 0$ must hold for every admissible set $(x,y,y',Y'') \neq (x,y,y',y'')$.

The proof given is not valid for an element(x,y,y',y'') at a corner point 3, but, since it holds at every point preceding 3, continuity considerations imply that it holds for the corner elements. Another corner condition can now be obtained. The first corner condition implies that

$$E(x,y,y'_-,y''_-,y'_+,y''_+) = -E(x,y,y'_+,y''_+,y'_-,y''_-)$$

at every point (x,y) of a maximizing arc. Since this E- function cannot be positive there follows the

SECOND CORNER CONDITION. At a corner (x,y) of a maximizing arc the condition

$$f(x,y,y'_-,y''_-) - y'_-f_{y'}(x,y,y'_-,y''_-) = f(x,y,y'_+,y''_+) - y'_+ f_{y'}(x,y,y'_+,y''_+)$$

must hold.

Since the maximizing arc satisfies $f_{y''}= 0$ a Taylor's expansion gives the equation

$$E(x,y,y',y'',Y',Y'') = \frac{1}{2}(Y'-y')^2 f_{y'y'} + (Y'-y')(Y''-y'')f_{y'y''} + \frac{1}{2}(Y''-y'')^2 f_{y''y''}$$

along E_{12} where the arguments of $f_{y'y'}$, $f_{y'y''}$, $f_{y''y''}$ are $[x,y,y'+\theta(Y'-y'),y''+\theta(Y''-y'')]$, $0 < \theta < 1$. It follows from

the Weierstrass necessary condition that this quadratic form in $Y'-y'$ and $Y''-y''$ must not be greater than zero for any set of values (x,y,y',y'',Y',Y'') for which (x,y,y',y'') is on E_{12} and (x,y,Y',Y'') is admissible.

ANALOGUE OF THE LEGENDRE NECESSARY CONDITION. At every element (x,y,y',y'') of a maximizing arc the conditions

$$f_{y''y''} \leqq 0, \qquad f^2_{y'y''} - f_{y'y'}f_{y''y''} \leqq 0, \qquad f_{y'y'} \leqq 0$$

must be satisfied.

6. Analogues of the Jacobi condition. The second variation of the integral I is defined to be the second derivative $I''(0)$ of the integral (4·5) taken at $a = 0$. The second variation of I taken over an extremal E_{12} is

$$I''(0) = \int_{x_1}^{x_2} 2\Omega(x,\eta,\eta',\eta'')dx + 2\left[f_y\eta + f_{y'}\eta'\right]^{x_2} \phi'(0)\alpha'(0)$$
$$+ \left[f\phi'' + \phi'^2\frac{d}{dx}f\right]^{x_2}\alpha'^2(0) + f\Big|^{x_2}\phi'(0)\alpha''(0),$$

where

$$2\Omega(x,\eta,\eta',\eta'') = f_{yy}\eta^2 + 2f_{yy'}\eta\eta' + f_{y'y'}\eta'^2 + 2f_{y'y''}\eta'\eta'' + f_{y''y''}\eta''^2 + 2f_{y''y}\eta''\eta$$

and the arguments of f and its partial derivatives are the elements $[x,y(x),y'(x),y''(x)]$ of the extremal E_{12}. If the curve N cuts E_{12} transversally but is not tangent to E_{12}, and if the values of $\alpha'(0)$ and $\alpha''(0)$ given by equations (4·3) and (4·4) are used the second variation takes the form

$$I''(0) = \int_{x_1}^{x_2} 2\Omega(x,\eta,\eta',\eta'')dx + A\,\eta^2(x_2),$$

where A is a constant defined by the equation

$$A = \left[\frac{2f_y\phi'}{\psi'-y'\phi'} + \frac{\phi'^2\frac{d}{dx}f}{(\psi'-y'\phi')^2} + \frac{f(y''\phi'^3-\phi'\psi''+\phi''\psi')}{(\psi'-y'\phi')^3}\right]_{\alpha=0}^{x=x_2}$$

A necessary condition for our problem is that the maximum of the second variation in the class of admissible variations $\eta(x)$, which are of class C' and satisfy the end condition $\eta(x_1) = 0$, must not be greater than zero. For the study of this maximum it is assumed that the arc E_{12} is of class C^{IV} and that the inequality $f_{y''y''}(x,y,y',y'') \neq 0$ holds along E_{12}. This problem is similar to the problem of maximizing the original integral I in the class of arcs $y = y(x)$ which are of class C' and have y'' continuous except at a finite number of places. The same kind of analysis that led to equation (4·8) yields the result that for every maximizing arc $\eta = \eta(x)$ in the present class of arcs there exist constants c_1 and c_2 such that the equation

$$\int_{x_1}^{x}\int_{x_1}^{x}\Omega_\eta dx - \int_{x_1}^{x}\Omega_{\eta'}dx + \Omega_{\eta''} = c_1(x-x_1) + c_2$$

is satisfied identically along the arc. It follows from this equation that every portion of a maximizing arc $\eta = \eta(x)$ which is of class C^{IV} must satisfy the equation[1]

$$\Omega_\eta - \frac{d}{dx}\Omega_{\eta'} + \frac{d^2}{dx^2}\Omega_{\eta''} = 0, \tag{6.1}$$

[1]This equation holds also when A = 0 and in this case is well known. See Hadamard, Calcul des Variations, § 124.

and at a point on a maximizing arc where η'' is discontinuous the corner conditions

$$(6\cdot2)\qquad \begin{aligned} &\Omega_{\eta''}\Big|_{-}^{+} = f_{y''y''}(\eta''_{+} - \eta''_{-}) = 0 \\ &\Omega_{\eta'} - \frac{d}{dx}\Omega_{\eta''}\Big|_{-}^{+} = -f_{y''y''}(\eta'''_{+} - \eta'''_{-}) - \left(\frac{d}{dx}f_{y''y''}\right)(\eta''_{+} - \eta''_{-}) = 0 \end{aligned}$$

must be satisfied. Since $f_{y''y''} \neq 0$ along E_{12} a consequence of these equations is that at every point on a maximizing arc where η' is continuous the derivatives η'' and η''' must also be continuous.

If in finding the maximum of the second variation the restriction that η' be continuous is removed the problem can be treated by the same kind of analysis that was used in §4 for the original problem. Results which are analogous to equations (4·11), (4·12), and the first corner condition are obtained and may be stated as follows. Between corners[1] of a maximizing arc $\eta = \eta(x)$ equations (6·1) and

$$(6\cdot3)\qquad \Omega_{\eta''} = f_{y''y''}\eta'' + f_{y'y''}\eta' + f_{y''y}\eta = 0$$

must be satisfied. At a corner the condition

$$(6\cdot4)\qquad \Omega_{\eta'}\Big|_{-}^{+} = f_{y'y''}(\eta''_{+} - \eta''_{-}) + f_{y'y'}(\eta'_{+} - \eta'_{-}) = 0$$

must hold and at the second end-point the transversality

---

[1]The word "corner" is used here in the same sense as in §4.

condition

$$(6\cdot5)\qquad \Omega_{\eta'}[x_2, \eta(x_2), \eta'(x_2), \eta''(x_2)] + A\,\eta(x_2) = 0$$

is necessary. Equations (6·3) and (6·4) imply that $f^2_{y'y''} - f_{y'y'}f_{y''y''} = 0$ at any place where either η' or η'' is discontinuous.

The necessary conditions derived above which a function $\eta(x)$ must satisfy if it maximizes the second variation may be summarized as follows: between corners the maximizing function $\eta(x)$ must satisfy equations (6·3) and

$$(6\cdot6)\qquad \Omega_\eta - \frac{d}{dx}\Omega_{\eta'} = 0.$$

At the second end-point the transversality condition (6·5) is necessary. The condition $f_{y''y''} \neq 0$ along E_{12} implies that η'' and η''' are continuous wherever η' is continuous. The inequality $\left[f^2_{y'y''} - f_{y'y'}f_{y''y''}\right]^{x_3} \neq 0$ implies that η' and η'' are continuous at x_3.

If the function $\eta(x)$ is of class C''' in the interval $x_4 \leqq x \leqq x_5$ the following equation holds.

$$(6\cdot7)\qquad \begin{aligned}\int_{x_4}^{x_5} 2\Omega\,dx &= \int_{x_4}^{x_5}\left[\Omega_\eta \eta + \Omega_{\eta'}\eta' + \Omega_{\eta''}\eta''\right]dx \\ &= \int_{x_4}^{x_5}\eta\left[\Omega_\eta - \frac{d}{dx}\Omega_{\eta'}\right]dx + \int_{x_4}^{x_5}\eta''\Omega_{\eta''}\,dx + \left[\eta\Omega_{\eta'}\right]_{x_4}^{x_5}\end{aligned}$$

If $\eta(x)$ is of class C^{IV} a further transformation can be made.

$$
(6\cdot 8)\qquad \int_{x_4}^{x_5} 2\Omega\, dx = \int_{x_4}^{x_5} \eta\left[\Omega_\eta - \frac{d}{dx}\Omega_{\eta'} + \frac{d^2}{dx^2}\Omega_{\eta''}\right] dx + \left[\eta\left(\Omega_{\eta'} - \frac{d}{dx}\Omega_{\eta''}\right)\right]_{x_4}^{x_5} + \left[\eta'\Omega_{\eta''}\right]_{x_4}^{x_5}
$$

FIRST ANALOGUE OF THE JACOBI CONDITION. <u>If</u> E_{12} <u>is an extremal of class</u> C^{IV} <u>with</u> $f_{y''y''} \neq 0$ <u>along it, and if</u> $u(x)$, <u>not identically zero, is a solution of class</u> C^{IV} <u>of equation</u> (6·1) <u>with either of the sets of properties</u>

$$
(a)\quad u(x_1) = u(x_3) = \left[u'\Omega_{\eta''}(x,u,u',u'')\right]^{x_1} = u'(x_3) = 0
$$

<u>or</u>

$$
(b)\quad u(x_1)=u(x_3)=\left[u'\Omega_{\eta''}(x,u,u',u'')\right]^{x_1} = \left[\Omega_{\eta''}(x,u,u',u'')\right]^{x_3} = 0,
$$

$$
\left[f^2_{y'y''} - f_{y'y'}\, f_{y''y''}\right]^{x_3} \neq 0,
$$

<u>where</u> $x_1 < x_3 < x_2$, <u>then</u> E_{12} <u>cannot furnish a maximum for the integral</u> I. <u>If these hypotheses hold except that</u> $x_3 = x_2$ <u>the arc</u> E_{12} <u>cannot furnish a maximum unless u satisfies the equation</u> $\Omega_{\eta''}(x,u,u',u'') = 0$ <u>and the end condition</u>

$\Omega_{\eta'}\left[x_2, u(x_2), u'(x_2), u''(x_2)\right] + Au(x_2) = 0.$

If such a function $u(x)$ exists the arc composed of $\eta = u(x)$ in the interval $x_1 \leqq x \leqq x_3$ and $\eta \equiv 0$ in $x_3 \leqq x \leqq x_2$ gives the expression $I''(0)$ the value zero as is easily shown by equation (6·8). However if $x_3 < x_2$ and u has

the properties (a) this arc cannot furnish a maximum for $I''(0)$ since, if it did, equations (6·2) imply that η'' and η''' must be continuous for $x = x_3$. Hence $u(x)$ satisfies the end condition $[u(x_3), u'(x_3), u''(x_3), u'''(x_3)] = [0,0,0,0]$. But the only solution of equation (6·1) which satisfies these initial conditions is $u \equiv 0$ which is contrary to hypothesis. Also if $x_3 < x_2$ and u has the properties (b) the composite arc cannot give $I''(0)$ a maximum value. For, if it did, equations (6·3) and (6·4) applied at $x = x_3$, together with the inequality $[f^2_{y'y''} - f_{y'y'}\, f_{y''y''}]^{x_3} \neq 0$ imply that $u'(x_3) = u''(x_3) = 0$. As before it follows that $u'''(x_3)$ is zero and hence $u \equiv 0$ which is contrary to hypothesis. The last statement of the condition follows from equations (6·3) and (6·5).

A similar proof establishes the

SECOND ANALOGUE OF THE JACOBI CONDITION. <u>Suppose E_{12} is an extremal of class C^{IV} which has $f_{y''y''} \neq 0$ along it and is not tangent to N at the point 2, and $u(x)$, not identically zero, is a solution of class C^{IV} of equation (6·1) with either of the sets of properties</u>

$$\text{(a)}\quad u(x_3) = \left[u\left(\Omega_{\eta'} - \frac{d}{dx}\Omega_{\eta''} + Au\right)\right]^{x_2} = \left[u \cdot \Omega_{\eta''}\right]^{x_2} = u'(x_3) = 0$$

<u>or</u>

$$\text{(b)}\quad u(x_3) = \left[u\left(\Omega_{\eta'} - \frac{d}{dx}\Omega_{\eta''} + Au\right)\right]^{x_2} = \left[u \cdot \Omega_{\eta''}\right]^{x_2} = \Omega_{\eta''}\Big|^{x_3} = 0,$$

$$\left[f^2_{y'y''} - f_{y'y'}\, f_{y''y''}\right]^{x_3} \neq 0,$$

where $x_1 < x_3 < x_2$ and the arguments of $\Omega_{\eta'}$ and $\Omega_{\eta''}$ in properties (a) and (b) are the sets (x,u,u',u''). Then E_{12} cannot furnish a maximum for the integral I. If these hypotheses hold except that $x_1 = x_3$ the arc E_{12} cannot furnish a maximum unless u satisfies the equation $\Omega_{\eta''}(x,u,u',u'') = 0$ and the end condition $\left[\Omega_{\eta'} + Au\right]^{x_2} = 0$.

The formula (6·7) is the basis for the

THIRD ANALOGUE OF THE JACOBI CONDITION. *Suppose E_{12} is an extremal with $f_{y''y''} \neq 0$ along it, and suppose $u(x)$, not identically zero, is a solution of class C''' of equation (6·3) such that $u(x_1) = u(x_3) = 0$, $\left[f^2_{y'y''} - f_{y'y'}\, f_{y''y''}\right]^{x_3} \neq 0$, and $\int_{x_1}^{x_3} u\left[\Omega_\eta - \frac{d}{dx}\Omega_{\eta'}\right] dx \geqq 0$ for $x_1 < x_3 < x_2$ where the arguments of Ω_η and $\Omega_{\eta'}$ in this integral are the sets (x,u,u',u''). Then E_{12} cannot furnish a maximum for the integral I. If these hypotheses hold except that $x_3 = x_2$ the arc E_{12} cannot furnish a maximum unless u satisfies the equation $\Omega_\eta(x,u,u',u'') - \frac{d}{dx}\Omega_{\eta'}(x,u,u',u'') = 0$.*

If such a function $u(x)$ exists the arc composed of $\eta = u(x)$ in the interval $x_1 \leqq x \leqq x_3$ and $\eta \equiv 0$ in $x_3 \leqq x \leqq x_2$ gives the integral $I''(0)$ a value greater than or equal to zero. However, if $x_3 < x_2$ this arc cannot furnish a maximum for $I''(0)$ since if it did equations (6·3) and (6·4) imply that $u'(x_3) = 0$. But the only solution of equation (6·3) having the initial values $[u(x_3), u'(x_3)] = [0,0]$ is $u \equiv 0$ which is contrary to hypothesis. The last statement of the condition follows from

the fact that a maximizing function $\eta(x)$ of the expression $I^{*}(0)$ must satisfy equation (6·6).

The value x_3 where $u(x)$ vanishes in the third analogue has an interesting geometric interpretation in case every semi-extremal is also an extremal. This hypothesis is quite restrictive but an example will be given in which it is satisfied. Consider a two-parameter family $y = y(x,a,b)$ of semi-extremals which includes E_{12} for $(a,b) = (0,0)$. Since $f_{y''}$ has continuous partial derivatives of the first three orders, the functions $y(x,a,b), y'(x,a,b), y''(x,a,b), y'''(x,a,b)$ are continuous with continuous partial derivatives.[1] If the function $y(x,a,b)$ is substituted in equation (5·2) an identity in x,a,b is obtained. Differentiation of this identity with respect to a and b respectively gives the equations

$$\Omega_{\eta''}(x, y_a, y'_a, y''_a) = f_{y''y''}\, y''_a + f_{y'y''}\, y'_a + f_{y''y}\, y_a = 0,$$

(6·9)

$$\Omega_{\eta''}(x, y_b, y'_b, y''_b) = f_{y''y''}\, y''_b + f_{y'y''}\, y'_b + f_{y''y}\, y_b = 0,$$

for $a = b = 0$. In these equations the arguments of the second partial derivatives of f are the elements belonging to E_{12} while the arguments of $y_a, y'_a, y''_a, y_b, y'_b, y''_b$ are $(x,0,0)$. Hence the functions $y_a(x,0,0)$, $y_b(x,0,0)$ are solutions of

---

[1]See Bliss, Princeton Colloquium Lectures, pp. 9-11, 95-98. The above statement is an immediate corollary of the theorems given by Bliss.

equation (6·3). It will be shown that, as a result of the non-vanishing of $f_{y''y''}$ on E_{12}, the family $y(x,a,b)$ can be so chosen that the determinant

$$\begin{vmatrix} y_a(x,0,0) & y_b(x,0,0) \\ y_a'(x,0,0) & y_b'(x,0,0) \end{vmatrix} \tag{6·10}$$

is different from zero at[1] $x = x_1$. Therefore the first row of this determinant is a fundamental system of solutions of the linear homogeneous differential equation (6·3), and since in this equation the coefficient $f_{y''y''}$ of y_a'' or y_b'' does not vanish in $x_1 \leqq x \leqq x_2$ then the determinant (6·10) does not vanish in $x_1 \leqq x \leqq x_2$.[2] Hence any solution of class C''' of equation (6·3) can be written in the form

$$u(x) = c_1 y_a(x,0,0) + c_2 y_b(x,0,0).$$

By the Sturm separation theorem any two solutions $u_1(x)$ and $u_2(x)$ of equation (6·3) which vanish for the same value of x are linearly dependent and have the same zeros.[3] The function

$$\Delta(x,x_1) = \begin{vmatrix} y_a(x,0,0) & y_b(x,0,0) \\ y_a(x_1,0,0) & y_b(x_1,0,0) \end{vmatrix}$$

is a solution of equation (6·3) which vanishes for $x = x_1$. Hence it has the same zeros as the function $u(x)$ of the third analogue. Furthermore, since the function $y(x,a,b)$ is also a

---

[1]See § 7, p. 46.

[2]Bolza, Variationsrechnung, p. 67.

[3]Bôcher, Methodes de Sturm (Paris, 1917), p. 46.

solution of equation (5·1) for every pair of values (a,b) it follows that $y_a(x,0,0), y_b(x,0,0)$ are also solutions of equation (6·6). This is proved by differentiation of equation (5·1) with respect to a and b. Hence in this case the condition $\int_{x_1}^{x_3} u(\Omega_\eta - \frac{d}{dx}\Omega_{\eta'})dx \geqq 0$ is always satisfied since the integrand function is identically zero. The zeros of the function $\Delta(x,x_1)$ different from x_1 define the points on E_{12} where it touches the envelope[1] of the one-parameter family of semi-extremals which pass through the point 1. A point at which E_{12} touches the envelope of this family of semi-extremals will be called a <u>conjugate point</u> of the point 1 on E_{12}. We have proved as a corollary of the third analogue the

FOURTH ANALOGUE OF THE JACOBI CONDITION. <u>If every semi-extremal is also an extremal, and if E_{12} is an extremal, with $f_{y''y''} \neq 0$ along it, which maximizes the integral I the point 1 can have no conjugate point interior to E_{12} at which $f^2_{y'y''} - f_{y'y'} f_{y''y''} \neq 0$.</u>

We shall later need the fact that if every semi-extremal is also an extremal, and if $f_{y''y''} \neq 0$ along E_{12} then the function $f^2_{y'y''} - f_{y'y'} f_{y''y''}$ is different from zero for every element on E_{12} or else it is identically zero along E_{12}. To prove this let $y = y(x,a,b)$ be a two-parameter family

---

[1]Bliss, Calculus of Variations (1925), p. 149.

of extremals which includes E_{12} for $(a,b) = (0,0)$. The equations

$$f_{y''y''}\, y_a''(x,0,0) + f_{y'y''} y_a'(x,0,0) + f_{y''y}\, y_a(x,0,0) = 0$$

$$(6\cdot 11)\quad f_{yy''}\, y_a''(x,0,0) + f_{yy'}\, y_a'\,(x,0,0) + f_{yy} y_a(x,0,0) -$$

$$\frac{d}{dx}\Big[f_{y'y''} y_a''\,(x,0,0) + f_{y'y'} y_a'\,(x,0,0) + f_{y'y}\, y_a(x,0,0)\Big] = 0$$

hold identically in x where the arguments of the partial derivatives of f are the elements $[x, y(x), y'(x), y''(x)]$ belonging to E_{12}. These equations are derived from equations (5·1) and (5·2) in the same way that equations (6·9) are found. If the functions $y_a''\,(x,0,0)$ and $y_a'''\,(x,0,0)$ are eliminated from equations (6·11) an equation of the form

$$B(x)\ y_a'(x,0,0) + C(x)\ y_a(x,0,0) = 0$$

is obtained. In a similar way we get the equation

$$B(x)\ y_b'(x,0,0) + C(x)\ y_b(x,0,0) = 0.$$

It has been remarked that the function $y(x,a,b)$ can be chosen so the determinant (6·10) does not vanish in the interval $x_1 \leqq x \leqq x_2$. It follows that $B(x) \equiv C(x) \equiv 0$. The function $B(x)$ can be written in either of the forms

$$B(x) = f_{y'y''} f_{y''y''}\,(2f'_{y'y''} + f_{y'y'}) - f^2_{y'y''}(f_{y'y''} + f'_{y''y''}) - f^2_{y''y''}\, f'_{y'y'}$$

$$= f_{y''y''}\, L'(x) - (f_{y'y''} + f'_{y''y''})L(x)$$

where $L(x) = f^2_{y'y''} - f_{y'y'} f_{y''y''}$. The function $C(x)$ is

$$C(x) = f_{y'y''} f_{y''y''} f'_{y''y} - f_{y'y''} f_{y''y} (f_{y'y''} + f'_{y''y''})$$
$$+ f_{y''y''} f_{y''y} (f'_{y'y''} + f_{y'y'} - f_{y''y}) + f^2_{y''y''} (f_{yy} - f'_{y'y}).$$

The equation $B(x) = 0$ implies

$$f^2_{y'y''} - f_{y'y'} f_{y''y''} = L(x) = ce^{\int p(x)\,dx}$$

where $p(x) = \dfrac{f_{y'y''} + f'_{y''y''}}{f_{y''y''}}$. Since $f_{y''y''} \neq 0$ in the interval $x_1 \leqq x \leqq x_2$, the function $L(x)$ is thus defined as a function which either does not vanish on the interval $x_1 \leqq x \leqq x_2$ or else is identically zero on this interval.

The proof of the fifth analogue of the Jacobi condition is analogous to the proof of the third.

FIFTH ANALOGUE OF THE JACOBI CONDITION. *Suppose* E_{12} *is an extremal which has* $f_{y''y''} \neq 0$ *along it and is not tangent to* N *at the point* 2, *and* $u(x)$, *not identically zero, is a solution of class* C''' *of equation* (6·3) *such that* $u(x_3) = 0$,

$$[f^2_{y'y''} - f_{y'y'} f_{y''y''}]^{x_3} \neq 0, \quad \int_{x_3}^{x_2} u\left(\Omega_\eta - \frac{d}{dx}\Omega_{\eta'}\right)dx \geqq 0, \text{ where}$$

$x_1 < x_3 < x_2$, *and either* (a) $u(x_2) = 0$ *or* (b) $\left[\Omega_{\eta'} + Au\right]^{x_2} = 0$, *where the arguments of* Ω_η *and* $\Omega_{\eta'}$ *are the sets* (x, u, u', u''). *Then* E_{12} *cannot furnish a maximum for the integral* I. *If these hypotheses hold except that* $x_1 = x_3$ *the arc* E_{12} *cannot furnish a maximum unless* u *satisfies the differential equation*

$$\Omega_\eta - \frac{d}{dx}\Omega_{\eta'} = 0.$$

If in this fifth condition the hypothesis that every semi-extremal is also an extremal is added, a new condition is obtained in the same way that the fourth was obtained from the third. This is actually two necessary conditions on the maximizing arc because of the alternate hypotheses (a) and (b). However in this case the condition with property (b') is stronger than that with property (a). For, let u_1 and u_2 be two solutions of class C''' of equation (6.3), neither of which is identically zero. Assume $u_1(x_3) = u_1(x_2) = \Omega_{\eta'}[x_2, u_2(x_2), u_2'(x_2), u_2''(x_2)] + Au_2(x_2) = 0$. The functions u_1 and u_2 are linearly independent, since if $c_1u_1 + c_2u_2 \equiv 0$, then $c_2u_2(x_2) = 0$. If $u_2(x_2) = 0$ and $[f^2_{y'y''} - f_{y'y'}f_{y''y''}]^{x_2} \neq 0$, the equations

$$\Omega_{\eta''}(x, u_2, u_2', u_2'')\Big|^{x_2} = f_{y''y''}(x_2)u_2''(x_2) + f_{y'y''}(x_2)u_2'(x_2) = 0$$

$$[\Omega_{\eta'}(x, u_2, u_2', u_2'') + Au_2]^{x_2} = f_{y'y''}(x_2)u_2''(x_2) + f_{y'y'}(x_2)u_2'(x_2) = 0$$

imply that $u_2'(x_2) = 0$; whence $u_2 \equiv 0$, contrary to hypothesis. Since u_1 and u_2 are linearly independent Sturm's separation theorem implies that u_2 vanishes for a value of x between x_3 and x_2. If every semi-extremal is an extremal then the function $f^2_{y'y''} - f_{y'y'}f_{y''y''}$ taken along E_{12} is non-vanishing or identically zero according as it is not zero or is zero at $x = x_2$. If it is identically zero the fifth condition cannot apply. Hence the sixth condition is not weakened if property

(a) is removed from it.

In order to state this condition geometrically consider a one-parameter family of semi-extremals of which E_{12} is a member, and a curve N which cuts each member of the family transversally. A point where E_{12} touches the envelope of this family of semi-extremals will be called <u>a focal point of the curve</u> N <u>on</u> E_{12}.

SIXTH ANALOGUE OF THE JACOBI CONDITION. <u>Suppose every semi-extremal is also an extremal, and</u> E_{12} <u>is an extremal, with</u> $f_{y''y''} \neq 0$ <u>along it, which is not tangent to the curve</u> N <u>at the point</u> 2. <u>Then if</u> E_{12} <u>maximizes the integral</u> I, <u>the curve</u> N <u>can have no focal point interior to</u> E_{12} <u>at which</u> $f^2_{y'y''}-f_{y'y'}f_{y''y''} \neq 0$.

We first show that, if every semi-extremal is an extremal, the value x_3 of the fifth condition is independent of the particular solution of equation (6·3) which is used to define it. It has been proved that under the assumptions of the sixth condition the inequality $[f^2_{y'y''} - f_{y'y'}f_{y''y''}]^{x_3} \neq 0$ for x_3 between x_1 and x_2 implies $[f^2_{y'y''} - f_{y'y'}f_{y''y''}]^{x_2} \neq 0$. Let u_1 and u_2, neither of which is identically zero, be two solutions of class C''' of equation (6·3) which satisfy the equations

$$\left[\Omega_{\eta'}(x,u_i,u_i',u_i'')+Au_i\right]^{x_2} = \left[f_{y'y''}u_i'' + f_{y'y'}u_i' + (f_{yy'}+A)u_i\right]^{x_2} = 0, \quad i=1,2.$$

By hypothesis

$$\Omega_{\eta''}(x,u_i,u_i',u_i'')\Big|^{x_2} = \left[f_{y''y''}u_i'' + f_{y'y''}u_i' + f_{yy'}u_i\right]^{x_2} = 0, \quad i = 1,2.$$

These four equations imply that the two linear homogeneous equations

$$f_{y'y''}(x_2)\,\zeta_1 + f_{y'y'}(x_2)\,\zeta_2 + (f_{yy'}(x_2) + A)\,\zeta_3 = 0$$

(6-12)

$$f_{y''y''}(x_2)\,\zeta_1 + f_{y'y''}(x_2)\,\zeta_2 + f_{y''y}(x_2)\,\zeta_3 = 0$$

in the three variables $(\zeta_1, \zeta_2, \zeta_3)$ have the two solutions $[u_1''(x_2), u_1'(x_2), u_1(x_2)]$ and $[u_2''(x_2), u_2'(x_2), u_2(x_2)]$. However the rank of the matrix of these equations is 2. Hence there cannot be two linearly independent solutions and the two solutions are proportional. Choose k so that $u_1(x_2) = ku_2(x_2)$. The value k is uniquely defined since $u_2(x_2) = 0$ and equations (6-12) would imply $u_2'(x_2) = 0$ and so $u_2 \equiv 0$, contrary to hypothesis. There follows the equation $u_1'(x_2) = ku_2'(x_2)$ and so $u_1 - ku_2$ is a solution of equation (6-3) which vanishes with its first derivative at $x = x_2$. Hence $u_1(x) \equiv ku_2(x)$ and the zeros of u_1 and u_2 are the same.

Next let $y = y(x,a)$ be a one-parameter family of extremals which includes E_{12} for $a = 0$ and such that each member is cut transversally by the curve N, that is, the equation

$$(6\text{-}13) \qquad \{f - y'[\phi(\alpha), a]\, f_{y'}\}\,\phi'(\alpha) + f_{y'}\,\psi'(\alpha) = 0,$$

where the arguments of f and $f_{y'}$ are the sets $\{\phi(\alpha), y[\phi(\alpha), a], y'[\phi(\alpha), a], y''[\phi(\alpha), a]\}$, becomes an identity in a when α is replaced by the function of a which is defined

by the system

$$(6{\cdot}14)\qquad y[\phi(\alpha), a] = \psi(\alpha), \qquad \alpha(0) = 0.$$

There exists a unique solution of this kind for values of a near 0 if E_{12} and N are not tangent at their point of intersection. If the two identities (6·13) and (6·14) are differentiated with respect to a, a taken equal to zero, and the quantity $\alpha'(0)$ eliminated between the two equations there results the equation

$$(6{\cdot}15)\qquad \Omega_{\eta'}[x_2, y_a(x_2,0), y_a'(x_2,0), y_a''(x_2,0)] + Ay_a(x_2,0)=0.$$

Since $y_a(x,0)$ is a solution of equations (6·3) and (6·6) and satisfies equation (6·15) the fifth condition prescribes that $y_a(x,0)$ must have no zeros in the interval $x_1 < x < x_2$. But this is equivalent to saying that the envelope of the family $y = y(x,a)$ must not touch E_{12} at an interior point.

The fourth analogue of the Jacobi condition now becomes superfluous since if an arc E_{12} satisfies the sixth analogue it necessarily satisfies the fourth. To prove this we first notice that, when every semi-extremal is an extremal, a function u(x) which is effective in the fifth condition cannot vanish at $x = x_2$, since, if it did, equations (6·12) would imply $u'(x_2) = 0$ and hence $u \equiv 0$, contrary to hypothesis. Therefore if E_{12} satisfies the sixth condition there can exist no such function which has a zero in the interval $x_1 < x \leqq x_2$. Then it follows from Sturm's separation theorem that any solution of equation (6·3) which vanishes at $x = x_1$ cannot have a

zero in the interval $x_1 < x \leqq x_2$. Hence the sixth condition implies the fourth condition in the strengthened form that allows the arc E_{12} to have no point, including the point 2, conjugate to the point 1. Furthermore in case there exists a two-parameter family of extremals it is the sixth analogue of the Jacobi condition which is useful in building a set of sufficient conditions as will be seen in §7.

7. Sufficient conditions. If every semi-extremal is also an extremal conditions can be given which are sufficient to insure that E_{12} makes the integral I a maximum. By a field of semi-extremals will be meant a region of the xy-plane associated with a one-parameter family $y = y(x,\alpha)$ of semi-extremals which simply covers it. The parameter value which gives the semi-extremal passing through the point (x,y) is $\alpha(x,y)$. The slope function is $p(x,y) = y'[x,\alpha(x,y)]$ and there is also associated with the field the function $r(x,y) = y''[x,\alpha(x,y)]$.

If $f_{y''y''} \neq 0$ along E_{12} equation (5·2) can be solved for y'',

$$y'' = G(x,y,y'), \tag{7·1}$$

and if $f_{y''}$ has continuous partial derivatives of the first n orders so has $G(x,y,y')$. Hence solutions of equation (7·1) which are of class C''' are semi-extremals. As for solutions of equation (7·1) it is known that through every element (x_3, y_3, y_3') in a neighborhood R' of E_{12} there passes a unique

solution $y = y(x,x_3,y_3,y_3')$ and that $y(x,x_3,y_3,y_3')$, $y'(x,x_3,y_3,y_3')$, and $y''(x,x_3,y_3,y_3')$ are continuous and have continuous partial derivatives of the first n orders.[1] Hence if E_{12} is a semi-extremal there exists a family of semi-extremals which includes E_{12} as a member. If R' is properly restricted, each of these solutions has an element corresponding to $x_3 = x_1$. The general solution then takes the form

$$y = y(x,y_3,y_3')$$

which is a two-parameter family including E_{12} for the parameter values $y_3 = y(x_1)$, $y_3' = y'(x_1)$. The equations expressing the fact that the above solution passes through (x_1,y_3,y_3') are $y_3 = y(x_1,y_3,y_3')$, $y_3' = y'(x_1,y_3,y_3')$. It follows from these equations that

$$\begin{vmatrix} y_{y_3}(x_1,y_3,y_3') & y_{y_3'}(x_1,y_3,y_3') \\ y'_{y_3}(x_1,y_3,y_3') & y'_{y_3'}(x_1,y_3,y_3') \end{vmatrix} = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1.$$

For convenience replace the parameter y_3 by a, and y_3' by b, and let (0,0) be the parameter values which give E_{12}.

An argument parallel to one used by Bliss now proves the[2]

LEMMA. <u>Every semi-extremal arc E_{12} having $f_{y''y''} \neq 0$ along it, and containing no point conjugate to 1, is interior</u>

---

[1]Bliss, Princeton Colloquium Lectures, pp. 95-98.
[2]Bliss, Calculus of Variations, pp. 155, 156.

to a field F of which it is itself a semi-extremal arc.

Assume that the arc E_{12} satisfies the hypotheses of the lemma and that every semi-extremal is also an extremal. Hence by the lemma a field F can be formed in which each of the semi-extremals is an extremal, and this field of extremals will imbed E_{12}. In such a field the differential of the integral I taken along any portion E_{34} of any of the extremals of the field such that E_{34} joins a curve C to a curve D can be found by a simple extension of the development of § 5. The expression for the differential is (5·4). Let 5 and 6 be the points on C and D respectively which are given by $b = b_5$. Then if both sides of (5·4) are integrated from $b = 0$ to $b = b_5$ we obtain

$$I(E_{56}) - I(E_{34}) = I^*(D_{46}) - I^*(C_{35}) \tag{7·2}$$

in which

$$I^*=\int\{f[x,y,p(x,y),r(x,y)]\,dx + (dy-p\,dx)f_{y'}[x,y,p(x,y),r(x,y)]\} .$$

In equation (7·2) if the curve D and the points 3,4,5,6 are fixed then everything is fixed except the path C_{35}. It follows that the integral I^* is independent of the path in the field F.

SUFFICIENT CONDITIONS FOR A STRONG RELATIVE MAXIMUM. Suppose that every semi-extremal is also an extremal, and that E_{12} is an admissible arc having the properties

a) it is an extremal having $f_{y''y''} < 0$ at every element

(x,y,y',y'') of E_{12},

b) it is cut transversally by N in a point which is not a singular point of N,

c) it contains no focal point of N,

d) $f[f^2_{y'y''} - f_{y'y'}f_{y''y''}] \neq 0$ on it at the point 2,

e) the inequality $f(x,y,y',y'')-f(x,y,p,r)-(y'-p)f_{y'}(x,y,p,r) < 0$ holds for every admissible element $(x,y,y',y'') \neq (x,y,p,r)$ where (x,y,p,r) are near the elements of E_{12}.

Then $I(E_{12})$ is greater than $I(C_{14})$ for every arc C_{14} distinct from E_{12} which joins the point 1 to the curve N and all of whose points (x,y) lie sufficiently near E_{12}.

In proving this theorem we first consider conditions under which a one-parameter family of extremals cut transversally by N and containing E_{12} can be constructed. Suppose N is represented by the equations $x = \phi(\alpha)$, $y = \psi(\alpha)$. By the previous theory of this section it follows that there exists a two-parameter family of extremals $y = y(x,a,b)$, which contains E_{12} for parameter values $(a,b) = (0,0)$, and having

$$(7\cdot3)\qquad \begin{vmatrix} y_a(x_2,0,0) & y_b(x_2,0,0) \\ y'_a(x_2,0,0) & y'_b(x_2,0,0) \end{vmatrix} \neq 0.$$

The equations

$$(7\cdot4)\qquad \begin{gathered} f[\phi(\alpha),\psi(\alpha),p,r]\phi'(\alpha)+[\psi'(\alpha)-p\,\phi'(\alpha)]\,f_{y'}[\phi(\alpha),\psi(\alpha),p,r] = 0 \\ y[\phi(\alpha),a,b] = \psi(\alpha), \quad y'[\phi(\alpha),a,b] = p, \quad y''[\phi(\alpha),a,b] = r \end{gathered}$$

are satisfied for $\alpha = 0$ by the values

$$(a,b,p,r) = (0,0,p_2,r_2) \tag{7·5}$$

of E_{12} at the point 2. If these equations can be solved for a,b,p,r as functions of α for values of α near 0 and such that for $\alpha = 0$ they assume the values (7·5) we will have a one-parameter family of extremals each of which is cut transversally by N. The functional determinant of equations (7·4) with respect to a,b,p,r at the point 2 given by $\alpha = 0$ is

$$\begin{vmatrix} 0 & y_a(x_2,0,0) & y'_a(x_2,0,0) & y''_a(x_2,0,0) \\ 0 & y_b(x_2,0,0) & y'_b(x_2,0,0) & y''_b(x_2,0,0) \\ [\psi'(0)-p_2\phi'(0)]\,f_{y'y'} & 0 & -1 & 0 \\ [\psi'(0)-p_2\phi'(0)]\,f_{y'y''} & 0 & 0 & -1 \end{vmatrix}$$

$$= [\psi'(0)-p_2\phi'(0)]\cdot\left[-f_{y'y'}\begin{vmatrix} y_a & y'_a \\ y_b & y'_b \end{vmatrix} -f_{y'y''}\begin{vmatrix} y_a & y''_a \\ y_b & y''_b \end{vmatrix}\right]$$

where the arguments of $f_{y'y'}$ and $f_{y'y''}$ are (x_2, y_2, p_2, r_2), and the arguments of y_a, y'_a, y''_a are $(x_2,0,0)$. If the equations (6·9) are solved for $y''_a(x_2,0,0)$ and $y''_b(x_2,0,0)$ and these values used in the above determinant it becomes

$$[\psi'(0)-p_2\phi'(0)]\cdot\left[-f_{y'y'} + \frac{f^2_{y'y''}}{f_{y''y''}}\right]\cdot\begin{vmatrix} y_a & y'_a \\ y_b & y'_b \end{vmatrix}. \tag{7·6}$$

This expression does not vanish if, (a) the point 2 is not a singular point on N, (b) $f \neq 0$ at the point 2,

(c) $f^2_{y'y''} - f_{y'y'}f_{y''y''} \neq 0$ at 2. For if $\psi' - p_2\phi' = 0$ it follows from the first of equations (7·4) and the condition (b) that $\phi'(0) = 0$. But $\phi'(0) = \psi'(0) - p_2\phi'(0) = 0$ implies that $\phi'(0) = \psi'(0) = 0$ at the point 2 which contradicts condition (a). Condition (c) implies that the second factor of (7·6) does not vanish and the last factor is different from zero because of the inequality (7·3).

Hence equations (7·4) can be solved for functions $a(\alpha)$, $b(\alpha)$, $p(\alpha)$, $r(\alpha)$ which become $0, 0, p_2, r_2$ for $\alpha = 0$. These functions are continuous and have continuous first derivatives near $\alpha = 0$. Thus each member of the one-parameter family of extremals $y = y(x,\alpha) = y[x, a(\alpha), b(\alpha)]$ is cut transversally by N. Since the equation $y[\phi(\alpha), \alpha] = \psi(\alpha)$ is an identity in α it may be differentiated with respect to α, giving $y_\alpha(x_2, 0) = \psi'(0) - p_2\phi'(0)$. This is not zero since it is the first factor of (7·6). If E_{12} contains no focal point of N, y_α remains different from zero along E_{12}. Hence E_{12} belongs to a field of extremals each member of which is cut transversally by N.

The integral I^* is independent of the path in this field and as a result of equation (4·12) equals zero when taken along N. Hence

$$I(C_{14}) - I(E_{12}) = I(C_{14}) - I^*(E_{12}) = I(C_{14}) - I^*(C_{14}) - I^*(N_{42})$$

$$= \int_{x_1}^{x_4} [f(x,y,y',y'') - f(x,y,p,r) - (y'-p)f_{y'}(x,y,p,r)]\,dx.$$

The sufficiency theorem follows immediately.

8. An example of the problem of § 4 for which a maximum exists. The second monopoly problem of § 1 which suggested the theory of §§ 4-7 can now be analyzed. The conclusions which were reached at the close of § 3 when we were considering the third monopoly problem can be easily generalized to fit the present problem. Since this is so and since we have no forms of the price and cost functions, (1·5) and (1·2), which have been shown to have special economic significance, we will examine instead another integral which satisfies the sufficient conditions of § 7.

If the function $f(x,y,y',y'')$ is defined by the equation

$$2f = a_{22} y''^2 + 2a_{21} y''y' + a_{11}y'^2 + 2a_{10} y'y + a_{00} y^2 + 2a_{02} yy'' + a_1 y' + a$$

where the coefficients a_{ij}, $(i,j = 0,1,2)$, a_1, and a are constants, and if the maximum of the integral $I = \int_{x_1}^{x_2} 2f\,dx$ is sought in the class of admissible arcs (in the sense of § 4) for which $y(x_1) = 0$ we have an example of the problem of § 4. Conditions on the coefficients of f will be sought which will insure that the integral I has a maximum in this class of arcs.

Hypotheses a) and d) of the sufficiency theorem and the analogue of the Legendre condition imply

$$(8\cdot1) \qquad a_{22} < 0, \qquad a_{11} < 0, \qquad a_{21}^2 - a_{11}a_{22} < 0.$$

The differential equations of the maximizing arc are

$$(8\cdot2) \qquad f_{y''} = a_{22} y'' + a_{21} y' + a_{02} y = 0$$

(8·3) $\frac{d}{dx} f_{y'} - f_y = a_{21} y''' + (a_{11} - a_{02}) y'' - a_{00} y = 0$

If y'' and y''' are eliminated from these two equations a third equation

(8·4) $a_{21}(a_{21}^2 - a_{11} a_{22}) y' + (a_{21}^2 a_{02} + a_{02}^2 a_{22} - a_{11} a_{02} a_{22} - a_{00} a_{22}^2) y = 0$

is obtained which the maximizing arc must satisfy. If every solution of equation (8·2) is to be a solution of equation (8·3) then equation (8·4) must be an identity in y and y'. This with the help of conditions (8·1) implies

(8·5) $a_{21} = 0 \qquad a_{00} = a_{02}(a_{02} - a_{11})/a_{22}$.

The form of the general solution of equation (8·2) depends upon the value a_{02}. Assume

(8·6) $a_{02} > 0$.

Then the general solution of equation (8·2) is

(8·7) $y = c_1 e^{mx} + c_2 e^{-mx}$

where $m = \sqrt{-a_{02}/a_{22}}$. The constants c_1 and c_2 are chosen to satisfy the end and transversality conditions.

(8·8)
$$y(x_1) = c_1 e^{mx_1} + c_2 e^{-mx_1} = 0$$
$$f_{y'}(x_2) = c_1(a_{11} m + a_{10}) e^{mx_2} + c_2(-a_{11} m + a_{10}) e^{-mx_2} + a_1/2 = 0$$

If

$$(8\cdot 9)\qquad \begin{vmatrix} e^{mx_1} & e^{-mx_1} \\ (a_{11}m + a_{10})e^{mx_2} & (-a_{11}m + a_{10})e^{-mx_2} \end{vmatrix}$$

$$= -2a_{11}m \cosh m(x_2 - x_1) - 2a_{10} \sinh m(x_2 - x_1) \neq 0$$

these equations have a unique solution for c_1 and c_2.

To obtain the focal points solve the second of equations (8·8) for one of c_1, c_2 in terms of the other. Since $a_{11}m \neq 0$ the coefficients $a_{11}m + a_{10}$ and $-a_{11}m + a_{10}$ cannot both vanish and this solution is always possible. Assume $-a_{11}m + a_{10} \neq 0$, then

$$c_2 = c_1 \frac{a_{11}m + a_{10}}{a_{11}m - a_{10}} e^{2mx_2} + \frac{a_1}{2(a_{11}m - a_{10})} e^{mx_2}.$$

Hence every member of the family

$$(8\cdot 10)\quad y = y(x, c_1) = c_1\left[e^{mx} + \frac{a_{11}m + a_{10}}{a_{11}m - a_{10}} e^{2mx_2} e^{-mx}\right] + \frac{a_1}{2(a_{11}m - a_{10})} e^{m(x_2 - x)}$$

is cut transversally by the line $x = x_2$. The abscissas of the points in which the maximizing arc touches the envelope of this one-parameter family of extremals are the zeros of the function

$$(8\cdot 11)\quad \begin{aligned} y_{c_1}(x, c_1) &= e^{mx} + \frac{a_{11}m + a_{10}}{a_{11}m - a_{10}} e^{2mx_2} e^{-mx} \\ &= \frac{2e^{mx_2} \cosh m(x_2 - x)}{a_{11}m - a_{10}} \left[a_{11}m + a_{10} \tanh m(x_2 - x)\right]. \end{aligned}$$

The fraction here cannot vanish. Any zeros the function may

have are zeros of the function

$$a_{11} m + a_{10} \tanh m(x_2 - x).$$

If $a_{10} \leqq 0$ this function is negative for all values of $x \leqq x_2$. If $a_{10} > 0$ and $\tanh m(x_2 - x_1) < -a_{11} m/a_{10}$ the function is non-zero for all values of x on the interval (x_1, x_2). These are necessary and sufficient conditions that the function (8·11) be non-vanishing in the interval $x_1 \leqq x \leqq x_2$. They also imply that the determinant (8·9) is not zero.

For every element (x,y,p,r) of the field of extremals which is cut transversally by the line $x = x_2$ the equation

$$2[f(x,y,y',y'') - f(x,y,p,r) - (y'-p)f_{y'}(x,y,p,r)]$$
$$= a_{22}(y'' - r)^2 + a_{11}(y' - p)^2$$

holds, where the element (x,y,y',y'') is any admissible element.

The only one of the sufficient conditions which remain to be verified is $f(x_2) \neq 0$. The constant $\underline{a}$ is not involved in the constants c_1, c_2, m and is independent of the other co-efficients of f. Hence in general $f(x_2)$ is not zero. Furthermore it is possible to simply cover the entire portion of the plane between $x = x_1$ and $x = x_2$ with the family of extremals (8·10).

THEOREM. <u>If</u> $a_{22} < 0$, $a_{11} < 0$, $a_{02} > 0$, $a_{21} = 0$, $a_{00} = \dfrac{a_{02}(a_{02} - a_{11})}{a_{22}}$, $f(x_2) \neq 0$, <u>and if either</u> $a_{10} \leqq 0$ <u>or</u> $\tanh\left[\sqrt{\dfrac{a_{02}}{-a_{22}}}(x_2 - x_1)\right] < \dfrac{-a_{11}}{a_{01}}\sqrt{\dfrac{a_{02}}{-a_{22}}}$ <u>then the arc</u> (8·7) <u>with the</u>

parameter values (c_1, c_2) which satisfy equations (8·8) furnishes a maximum for the integral I among all admissible arcs which join the point $(x_1, 0)$ to the line $x = x_2$.

FUNCTIONS OF LINES AND THE CALCULUS OF VARIATIONS

BY

RALPH GRAFTON SANGER

C O N T E N T S

INTRODUCTION

The purpose of this dissertation is to show the relations existing between the theory of functions of lines and the calculus of variations, and to trace the historical development of these relationships. In order to do this it has been found necessary to study first the various definitions of the differentials of a function of a line, and second, the applications of such differentials to the problems of the calculus of variations. In the following pages the terms "function of a line" and "functional" will be used interchangeably.

In 1887 the first precise definition of a functional was made by Volterra. Between 1887 and 1890 several papers by Volterra on this subject appeared in the Rendiconti dei Lincei. From then until about 1913 little progress was made. The renaissance of the theory of functions of lines may be said to begin with the publication in 1913 of two books by Volterra; one on functions of lines and the other on integral equations. About the same time there appeared some articles by Fréchet on differentials of functions and functionals, which also stimulated interest in the field. Since then

many articles have appeared on various phases of the subject. In this paper only those articles which deal with the derivatives or differentials of functionals, or phases of the theory of functionals which may be applied to problems of the calculus of variations, are considered.

In the first chapter the definition of a functional as conceived by Volterra, together with the study of the derivatives and differentials of a functional as deduced by him are presented. As Volterra was the first person to consider such ideas, his work is first discussed. The criticisms of his derivatives and differentials by Hadamard and Bliss, showing that they must be modified in order to be applicable to problems such as those of the calculus of variations, are also considered. Fischer has extended the definition of a derivative of a functional as given by Volterra so that it is applicable to the problems of the calculus of variations. He also generalized these ideas for functions of surfaces. The topics indicated above are in essence, the content of the first chapter.

The second chapter deals primarily with the conception of differentials of functionals as given by Fréchet, together with the modifications of his definitions which were made by Le Stourgeon, and with theorems concerning the representation of linear functionals by means of Stieltjes integrals. The definitions of the differentials of a functional as proposed by Fréchet appear to be inconvenient in the theory of the calculus of variations. Le Stourgeon has

modified Fréchet's definitions in such a manner that they are applicable to this theory. According to Fréchet's definition the first differential of a functional is a linear functional, and can therefore be represented as a Stieltjes integral. A brief discussion of such representations of linear and bilinear functionals, in particular of the Riesz theorem, is also made in this chapter, as such representations are useful in the applications of the theory of functionals to the calculus of variations.

A study of the applications of the theory of functionals to the problems of the calculus of variations is the theme of chapter three. The more or less unsuccessful attempts to relate the two fields, such as were made by Hadamard and Volterra are first considered. The first systematic study of the problem of minimizing a functional was made by Le Stourgeon. Her work and its applications to the simple problem of the calculus of variations are treated in some detail. The work of Fischer on the same problem is also presented. Fischer further studied the problem of minimizing a function of a surface, a generalization of the problem of minimizing a double integral. The problem of Lagrange has been treated from the standpoint of functions of lines, by Hahn and others, and a section is devoted to such treatments. The Hamilton-Jacobi theory for double integrals and its relations to the theory of functions of lines are also included in this chapter. The relations of the theory of functionals to the theory of the calculus of variations as expounded by Tonelli

and his followers are likewise briefly considered.

The theory of the calculus of variations is not usually thought of as a chapter in the theory of functionals. The functionals of the calculus of variations are relatively so special that many of their most important properties do not hold for functionals in general. It is probably for this reason that the contributions of the general theory to the theory of the calculus of variations are so few in number. The purpose of this paper is to make a survey of the relationships hitherto developed between the general and the more special theories, with the hope that further results of interest in this domain may be secured in the near future.

CHAPTER I

VOLTERRA'S DERIVATIVE OF A FUNCTIONAL

Introduction. It is seldom that in any mathematical field one can point to a certain article and say that with this article the study of that particular field begins. For the case of the theory of functions of lines, however, the theory may be said to begin quite definitely with an article by Volterra written in 1887 [2, p. 97] *. Previous writers had studied functions which depended on other functions, as for example Koenigs in his work on functional equations [1, p. S. 3], but they did not consider functions which depended on all the values of another function, and they consequently did not have need to define a functional of the most general type.

In this first chapter of the present paper the work of Volterra, with its criticisms and generalizations by others, is presented.

1. The Volterra derivative. Volterra [2, p. 97] defined a function of a line, or functional, as a variable F whose value was determined when all the values of a function y(x) defined on a certain interval ab were given, and denoted his functional by the symbol

$$F = F|[y\overset{b}{\underset{a}{(x)}}]|.$$

* The symbol [2, p. 97] means reference 2 of the bibliography at the end of the text, page 97. Other symbols of this type should be similarly interpreted.

If the functional depended also on a variable t, it was expressed by the notation

$$F = F\left|\left[y\left(\overset{t}{\underset{a}{x}}\right), t\right]\right|.$$

Similarly, functionals which depended on n functions and m variables were represented in the form

$$F = F\left|\left[y_1\left(\overset{t_1}{\underset{a_1}{x}}\right), \cdots y_n\left(\overset{t_n}{\underset{a_n}{x}}\right), t_1, t_2, \cdots, t_n\right]\right|.$$

For such a function of a line F $\left|\left[y(x)\right]\right|$ Volterra defined a derivative at a point x = ξ. [2, p. 97; 3, p. 141; 17, p. 24; 18, p. 12; 38, p. 23]. For the definition of the derivative the functional F is assumed to have continuity of order zero, that is, for every $\varepsilon > 0$ it is true that there exists a δ such that when $|\eta(x)| < \delta$ on ab, then $\left|F\left|\left[y(x)+\eta(x)\right]\right| - F\left|\left[y(x)\right]\right|\right| < \varepsilon$. The functions y(x) on which the functional depends are assumed to be continuous on the interval ab. On a subinterval h of ab containing the point ξ let y(x) have a continuous increment $\eta(x)$ such that $|\eta(x)| < \varepsilon$. The curve y = y(x) so modified will be represented by the equation y = y(x) + η (x) whose graph is shown in the figure. In order to define a derivative, the following further relations are supposed to be satisfied.

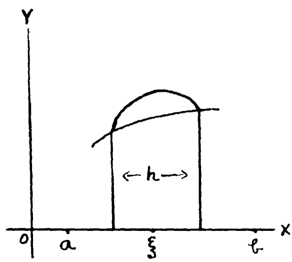


Fig. 1.

I. The variations η of the ordinates are less in

absolute value than ε , and for the corresponding increment of F, represented by Δ F, the inequality

$$\left| \frac{\Delta F}{\varepsilon h} \right| < M,$$

is presupposed where M is a constant.

II. If the deformation of the curve y = y(x) on the interval h is always on the same side of the curve, that is, obtained by either increasing or decreasing all the ordinates, and if σ represents the area between the original and the varied curve, then Volterra assumes that the limit

$$\lim_{\varepsilon, h \doteq 0} \frac{\Delta F}{\sigma}$$

exists. This limit is a function of the curve y = y(x) and the point ξ into which the interval h shrinks, and will be denoted by $F' | [y \overset{t}{\underset{a}{(x)}}, \xi] |$.

III. With respect to all functions y(x) and all abscissae ξ the ratio $\Delta F / \sigma$ tends uniformly towards its limit.

IV. The functional $F' | [y \overset{t}{\underset{a}{(x)}}, \xi] |$ is continuous with respect to y(x) and ξ.

The quantity $F' | [y \overset{t}{\underset{a}{(x)}}, \xi] |$ is then called the derivative of the functional F $| [y(x)] |$ taken at the point x = ξ .

One should note that the derivative so defined is applicable only to functionals which have continuity of order

zero. The assumption III, that the limit of $\Delta F/\sigma$ is uniformly approached as ε and h tend toward zero is a very strong one.

The problems of the calculus of variations are problems in the theory of functions of lines for which F has the form

$$F\,|[y(x)]| = \int_a^b f(x, y(x), y'(x))\, dx.$$

These functionals obviously do not in general have continuity of order zero, and therefore do not always have derivatives in the sense described above. The derivative as defined above is taken at the particular point on the arc y(x) corresponding to $x = \xi$. For the problems of the calculus of variations it turns out that a differential defined along y(x) over the whole interval $a \leqq x \leqq b$ has greater applicability.

2. <u>The Volterra differential.</u> The differential of a functional, as contrasted to its derivative, was defined by Volterra [2, p. 101] and shown to be expressible in terms of the derivative in the form

(2.1) $$\delta F = \int_a^b F'\,|[y(x), \xi]|\, \delta y(\xi)\, d\xi.$$

In his first paper on functions of lines [2, pp. 100-2] Volterra gives a proof of this formula which is not altogether satisfactory. In his later works [17, p. 25; 18, p. 22; 38, p. 23] he obtains the formula for the differential by analogy with the case for a function of a real variable. In that case, if F is a function of n real variables y_i and

if δy_i is the differential of y_i, the differential of F is a linear expression with suitable properties of the form

$$\delta F = \sum_1^m F_i' \delta y_i,$$

where the symbols F_i' designate simply the coefficients of the differentials δy_i without any reference to their interpretations as derivatives. In the case of functions of lines one has an analogous formula (2.1), the summation sign being replaced by an integral. This method of attack provides the definition of a differential without the preliminary definition of a derivative.

In his first paper [2, p. 97] on the derivative of a functional Volterra assumed that the ratio $\Delta F/\varepsilon h$ (see page 7) was bounded. In a second article [3, p. 141] he considered cases when this ratio had other particular properties. In one of them he showed that if the variation η of y(x) was less than ε on the interval h, and at the point ξ the relation

$$\lim_{\varepsilon, h \doteq 0} \frac{\Delta F}{\varepsilon} = 0,$$

holds, then the differential of F still has the form

$$\delta F = \int_a^b F'|[y(x), t]|\,\eta(t)\,dt.$$

Another case which he discussed [3, p. 145] was one for which the functions y(x) and their variations η had

continuous derivatives up to and including those of order m. Let y be given a variation η on the interval h containing the point ξ, and let η be so chosen that $\eta, \eta', \cdots \eta^{(m)}$ are in absolute value less than ε and respectively equal to $\rho_0, \rho_1, \cdots, \rho_m$ at the point ξ. Furthermore it is presupposed that the ratios ρ_i/ε tend towards a set of constants k_i as ε and h approach zero. For every such approach to zero it is also assumed that at the point ξ the ratio $\Delta F/\varepsilon$ has the limit

$$\lim_{\varepsilon, h \doteq 0} \frac{\Delta F}{\varepsilon} = \sum_0^m a_p K_p,$$

where the coefficients a_p have well determined finite values. Under these circumstances Volterra shows that the differential of F has the form

$$\delta F = \int_a^b F'|[y(x), t]|\eta(t)\,dt + \sum_0^m a_p \eta^{(p)}(\xi).$$

In order to prove this, let us define a functional G by means of the relation

$$G|[y(x)]| = F|[y(x)]| - \sum_0^m a_p y^p(\xi). \tag{2.2}$$

From this definition it follows with the help of the proof first mentioned in this section that G has a differential in the form

$$\delta G = \int_a^b G'|[y(x), t]|\eta(t)\,dt,$$

and it may be easily proved that

$$G' \left| \left[y(x), t \right] \right| = F' \left| \left[y(x), t \right] \right| .$$

Thus, taking the differential of both sides of equation (2.2) and solving for δF, it is clear that

$$\begin{aligned} \delta F &= \delta G + \sum_{0}^{m} a_p \delta y^{p}(\xi) \\ &= \int_a^t F' \left| \left[y(x), t \right] \right| \eta(t)\, dt + \sum_{0}^{m} a_p \eta^{(p)}(\xi) . \end{aligned}$$

Such an expression for the differential is interesting, but does not appear to be applicable to many particular problems.

Volterra defines second derivatives of functionals by a method similar to the one used in deducing first derivatives. The first derivative is defined at a point ξ_1, and thereafter the second derivative at a point ξ_2, and the latter may be expressed in the form

$$F'' \left| \left[y\overset{t}{\underset{a}{(x)}}, \xi_1, \xi_2 \right] \right| .$$

Under suitable hypotheses this expression is proved to be symmetric in ξ_1 and ξ_2 [2, p. 102; 17, p. 26]. Volterra also shows that for $F \left| \left[y(x) + \varepsilon \eta \right] \right|$ the derivative with respect to ε at $\varepsilon = 0$ is expressible in the form

$$\left(\frac{dF}{d\varepsilon} \right)_0 = \int_a^t F' \left| \left[y(x), \xi \right] \right| \eta(\xi)\, d\xi ,$$

and in an analogous manner the second derivative may be represented in the form

$$\left(\frac{d^2 F}{d\varepsilon^2} \right)_0 = \int_a^t \eta(\xi_1)\, d\xi_1 \int_a^t F'' \left| \left[y(x), \xi_1, \xi_2 \right] \right| \eta(\xi_2)\, d\xi_2 .$$

These expressions correspond to the first and second differentials used in the theory of the calculus of variations.

3. The criticism of the derivative by Bliss. The Volterra derivative as defined on page 7 has been shown not to be applicable to problems of the calculus of variations. As such examples constitute an important sub-field of the theory of functionals, it is desirable that there be a theory suited to such applications. Bliss [23, p. 175] has given examples from the calculus of variations for which the Volterra derivative does not exist.

One example for which the derivative does not exist, because the integral is not continuous in the Volterra sense, is the length integral

$$F = \int_a^b \sqrt{1+y'^2}\, dx .$$

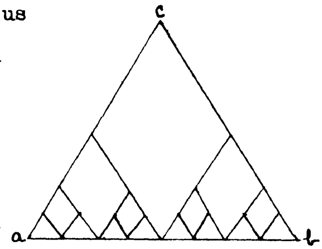


Fig. 2.

In the figure, ac + cb is the length of each of the broken lines connecting a and b, which consist of the slanting sides of the triangles with bases on ab and equal altitudes. In each neighborhood of ab there is one of these paths, all of the same length, and one readily sees by means of these variations of the straight line ab that the length integral has not continuity of order zero as presupposed by Volterra.

Another example is given where the limits $\Delta F/\sigma$ and $\Delta F/\epsilon h$ (see page 7) are not well defined as required by Volterra. Consider an integral of the form

$$F = \int_a^b f[x, y(x), y'(x)]\,dx, \tag{3.1}$$

where the function f is supposed to be continuous and to have continuous first and second derivatives in a neighborhood R of the values (x, y, y') belonging to the curve $y = y(x)$, $(a \leq x \leq b)$. If $\eta(x)$ is the variation of y, the corresponding increment ΔF of the integral (3.1) may be expressed in the form

$$\begin{aligned} \Delta F &= \int_a^b \{f(x, y+\eta, y'+\eta') - f(x, y, y')\}\,dx \\ &= \int_a^b \{A\eta + B\eta'\}\,dx, \end{aligned} \tag{3.2}$$

* where

$$A = \int_0^1 f_y(x, y+\theta\eta, y'+\theta\eta')\,d\theta,$$

$$B = \int_0^1 f_{y'}(x, y+\theta\eta, y'+\theta\eta')\,d\theta,$$

and where η is continuous and has continuous first and second derivatives except perhaps at a finite number of points on the interval ab. The values $(x, y+\eta, y'+\eta')$ for $a \leq x \leq b$ are assumed to be in R. After an integration by parts, usual in the calculus of variations, and an application of the mean value theorem the expression (3.2) becomes

$$\Delta F = \int_a^b \left(A - \frac{dB}{dx}\right)\eta\,dx = \left[A - \frac{dB}{dx}\right]_{x=x'} \int_a^b \eta\,dx, \tag{3.3}$$

where x' is a suitably chosen point on the interval of length h containing the point ξ at which the Volterra derivative

was taken. If the variation η (x) has the form

$$\eta(x) = \sqrt{r^2-(x-\xi)^2} - (r-\varepsilon),$$

corresponding to a circular arc on the interval

$$\xi - \frac{h}{2} \leq x \leq \xi + \frac{h}{2},$$

where it is not identically zero, as shown in figure 3, then one sees that

$$\sigma = \int_a^b \eta' dx = r^2(\theta - \sin\theta\cos\theta),$$

$$(3.4)$$

$$\varepsilon h = 2r^2 \sin\theta(1-\cos\theta)$$

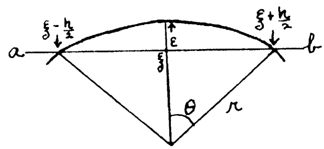


Fig. 3.

When ε is allowed to approach zero while r remains constant, the above relations imply that h and θ both approach zero and

$$\lim_{\varepsilon \doteq 0} \frac{\Delta F}{\sigma} = \left[f_y - \frac{d}{dx} f_{y'} + \frac{1}{2} f_{y'y'} \frac{1}{r}\right]_{x=\xi},$$

$$\lim_{\varepsilon \doteq 0} \frac{\Delta F}{\varepsilon h} = \frac{4}{3}\left[f_y - \frac{d}{dx} f_{y'} + \frac{1}{2} f_{y'y'} \frac{1}{r}\right]_{x=\xi}.$$

At a point ξ where $f_{y'y'}$ is not zero either one of the last two limits may be made to take on an arbitrarily assigned value by properly choosing r. Hence here the condition that the above two ratios are uniquely defined or bounded is not satisfied.

For a general variation η satisfying restrictions

of the type imposed in the last paragraph it follows from (3.3) and (3.4) that

$$\frac{\Delta F}{\sigma} = \Big[\int_0^1 \{f_y - f_{y'x} - f_{y'y}\,y' - f_{y'y'}\,y''\}\,d\theta$$
$$-\eta'\int_0^1 f_{y'y}\,\theta\,d\theta - \eta''\int_0^1 f_{y'y'}\,\theta\,d\theta\Big]_{x=x'},$$

where the arguments of f and its derivatives are $(x, y + \theta\eta, y' + \theta\eta')$ and x' is the intermediate value between $\xi - \frac{h}{2}$ and $\xi + \frac{h}{2}$ mentioned on the preceding page. If this limit is to exist and have the value

$$\lim_{\varepsilon, h \doteq 0} \frac{\Delta F}{\sigma} = \Big[f_y - \frac{d}{dx} f_{y'}\Big]_{x=\xi},$$

it is evident that the conditions

$$|\eta(x)| < \varepsilon, \quad |\eta'(x)| < \varepsilon, \quad |\eta''(x)| < \varepsilon, \quad (a \leq x \leq b)$$

should also be imposed on η. If it is also further presupposed that

$$\Big|f_y - \frac{d}{dx} f_{y'}\Big| < M, \quad |f_{y'y}| < M, \quad |f_{y'y'}| < M, \quad (a \leq x \leq b)$$

where M is a fixed quantity, then one is assured that the ratio $\Delta F/\varepsilon h$ is bounded since

$$\left|\frac{\Delta F}{\varepsilon h}\right| = \left|\frac{\Delta F}{\sigma}\right| \cdot \left|\frac{\sigma}{\varepsilon h}\right| \leq \left|\frac{\Delta F}{\sigma}\right| \leq M + M\frac{\varepsilon}{2} + M\frac{\varepsilon}{2}.$$

Bliss also shows what properties possessed by an integral of the form

$$F[y] = \int_a^b f(x, y, y', \cdots, y^{(n)})\, dx,$$

will insure that it has a derivative in the Volterra sense approached with order r. A more detailed study of this work was undertaken by Fischer, and will be considered in Section 7.

From the considerations above, it is seen that the Volterra derivative has to be modified in order to be applicable to the problems of the calculus of variations, and the type of modification is indicated in the preceding paragraphs.

4. Volterra's work on the three-space problem. In a series of articles Volterra developed a theory of functionals [2, p. 97; 3, p. 141; 4, p. 153] and applied his idea of a functional derivative to various examples, such as functionals in three-space [5, p. 225; 6, p. 274], ordinary monogenic functions [7, p. 281; 8, p. 107; 9, p. 196; 13, p. 233], monogenic functions in higher dimensional spaces [10, p. 158; 11, p. 291; 12, p. 599], and the Hamilton-Jacobi theory of the calculus of variations for multiple integrals [14, p. 127]. In his texts [17, p. 77; 18, p. 29; 38, p. 41] he presents the theory of functions of lines as preliminary to a study of integral and integro-differential equations. For some of the problems of the calculus of variations it has been shown that the Volterra derivative will not always exist. However, Volterra's applications to the calculus of variations, and to the theory of maxima and minima of more general functionals than those which

occur in the calculus of variations have been very suggestive.

In particular, for the space case [5, p. 226] for which a curve L is defined by the equations

$$x = x(s), \qquad y = y(s), \qquad z = z(s),$$

and a neighboring curve L_1 by the equations

$$x = x_1(s), \qquad y = y_1(s), \qquad z = z_1(s),$$

the variations of the curve L are the quantities

$$\delta x = x_1(s) - x(s), \quad \delta y = y_1(s) - y(s), \quad \delta z = z_1(s) - z(s).$$

Volterra assumed that the differential of a functional $F|[L]|$ had the form

$$\delta F = \int_L (F_x' \delta x + F_y' \delta y + F_z' \delta z) ds.$$

This is a direct extension of the differential of a functional depending on one independent variable.

Volterra [6, p. 274] also considers the differential of a functional in three space for certain special cases. In each instance a formula for the differential is deduced. For example, if α, β, γ are the direction cosines of the tangent to the curve L, then it turns out that the relation $\alpha F_x' + \beta F_y' + \gamma F_z' = 0,$

is satisfied, and the quantities F_x', F_y', F_z' may be expressed in the form

$$F_x' = \gamma B - \beta C, \qquad F_y' = \alpha C - \gamma A, \qquad F_z' = \beta A - \alpha B.$$

The differential may then be represented by the expression

$$\delta F = \int_L \left\{ A(\beta\delta z - \gamma\delta y) + B(\gamma\delta x - \alpha\delta z) + C(\alpha\delta y - \beta\delta x)\right\} ds. \tag{4.1}$$

Prange, in his dissertation [30, p. 41] calls the quantities A, B, C, which occur in equation (4.1) the derivatives of the functional F with respect to the coordinate planes. He makes use of them in the Hamilton-Jacobi theory for double integrals. Such applications will be considered in a later section.

Fabri [15, p. 432] has generalized the ideas of Volterra to functionals of the type

$$F = F\left|\left[y(x_1, x_2, \cdots, x_n)\right]\right|,$$

where the x_i are n independent variables ranging over a domain C. The first variation of such a functional is found in the form

$$\frac{dF}{d\varepsilon} = \int_C \eta(\xi_1, \cdots, \xi_n) F'\left|\left[y(x_1, \cdots, x_n)\, \xi_1, \cdots, \xi_n\right]\right| d\xi_1 \cdots d\xi_n,$$

where
$$F' = \lim_{\varepsilon, C_1 \doteq 0} \frac{F\left|[y+\varepsilon\eta]\right| - F\left|[y]\right|}{\int_{C_1} \eta\, dx_1 \cdots dx_n}$$

η being the customary variation of y different from zero on the portion C_1 of C. Fabri also considered functionals in three-space, and deduced representations of special functionals and their differentials in forms similar to those obtained by Volterra.

The work of Volterra and Fabri for the three-space problem is exceedingly suggestive of possible applications of the theory of functionals. Their results, however, need simplification and adaptation.

5. Hadamard's criticism. A criticism of the Volterra derivative of a function of a line L in three-dimensional space was also made by Hadamard [16, p. 40]. He noted without proof for a function of the form

$$F|[L]| = \int_L f\left(x, y, z, \frac{dy}{dx}, \frac{dz}{dx}, \frac{d^2y}{dx^2}, \frac{d^2z}{dx^2}, \cdots\right) dx,$$

that under certain circumstances the expression for the differential

$$\delta F = \int_L (F_x' \delta x + F_y' \delta y + F_z' \delta z)\, ds, \tag{5.1}$$

as given by Volterra [5, p. 227] was not valid unless one added to it terms of the form

$$P_x \delta x + P_y \delta y + P_z \delta z + P_{y'} \delta y' + P_{z'} \delta z' + \cdots .$$

Similarly it was seen that the differential of the derivative F_x' could not be represented in the form (5.1), as it

depended on the quantities y' and z' as well as x,y,z.

Hadamard noted that the differential of Volterra was a distributive function as defined by Pincherle, an additive function as defined by Bourlet, or a linear function in δx, δy, δz according to his own definition. He also showed that many of the properties of the Volterra derivative for the functional F may be obtained by considering the values of the functional F taken along a two-parameter family of curves L, depending on the parameters λ and μ. The identity

$$\frac{\partial^2 F}{\partial\lambda\partial\mu} = \frac{\partial^2 F}{\partial\mu\partial\lambda},$$

is then satisfied. Hadamard states that from this formula alone many of the relations involving the inversion of various differentiations of a functional may be obtained.

Hadamard's criticism was the first to be made of the Volterra differential. He showed that such differentials do not necessarily exist for all types of functionals, and indicated the types of functionals for which they were applicable. He did not carry through the details of the suggestions which he made.

6. <u>Other hypotheses for the existence of the Volterra derivative.</u> Other hypotheses for the existence of the Volterra derivative of a function of a line have been given by Evans. His purpose was to weaken the continuity properties required for the functional $F\Big|\big[y(x)\big]\Big|$, and also for

its argument $y(x)$, in order that the derivative should exist.

One set of hypotheses given by Evans [28, p. 387] consists of the following three assumptions:

I. The function $y(x)$ lies between two functions $G(x)$ and $H(x)$ such that $G(x) < H(x)$ on the interval ab, and $F|[y]|$ is well defined and continuous for every continuous function $y(x)$ in that region.

II. On the interval h, within the interval ab, let $y(x)$ take on a continuous variation $\eta(x)$, vanishing at the ends of h, of one sign, and such that $|\eta(x)| < \varepsilon$ in the interior of h. Let the corresponding increment of F be denoted by ΔF, and for convenience, write

$$\sigma = \int_h \eta \, dx .$$

It is presupposed that the ratio $\Delta F/\sigma$ approaches a fixed limit F' as ε and h tend towards zero, provided that the interval always contains the point ξ at which the derivative is to be taken.

III. The ratio $\Delta F/\sigma$ approaches its limit uniformly with respect to all the functions $y(x)$ described above and all points ξ on the interval ab.

If these conditions are fulfilled Evans shows that the derivative of a functional will surely exist. These assumptions are similar to those made by Volterra, the differences being in condition I and the fact that the derivative is not assumed to be continuous. Evans gives two other sets of hypotheses which assure the existence of the

Volterra derivative.

Daniell [35, p. 414] deduced a criterion for the linear property of the Volterra differential. He considered only continuous functions y(x) and defined the differential by the relation

$$D(y,\eta) = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\left\{F|[y+\varepsilon\eta]| - F|[y]|\right\}.$$

In place of Volterra's four conditions he made the following two assumptions:

I. $F|[y]|$ satisfies a Cauchy-Lipschitz condition, which states that there is a number M such that

$$\left|F|[y_1]| - F|[y_2]|\right| \leq M \cdot \max|y_1(x) - y_2(x)|.$$

II. The differential $D(y_1,\eta)$ exists for all continuous η, and all continuous y_1 in the neighborhood of y.

Under these conditions Daniell shows that the differential $D(y,\eta)$ is a linear functional.

These articles by Evans and Daniell are of interest because they show that the hypotheses Volterra made for the existence of the derivative and the differential of a functional are unnecessarily stringent.

7. <u>The generalized Volterra derivative of Fischer.</u> It has been shown that the derivative of a functional as conceived by Volterra frequently needs modification in order to be applicable to the problems of the calculus of variations. The purpose of Fischer was to extend the

Volterra derivative in such a way that it may be utilized in the theory of the problem of Lagrange in the calculus of variations.

Fischer [19, p. 375] defined a derivative of a function of a line L(C), relative to a specially defined class of curves K in the plane and a set of points $\xi_1, \cdots, \xi_m$ on the interval x_1x_2. In order to define the set of curves K, consider a set of differential equations

$$z_i' = g_i(x, y, y', z_1, \cdots, z_m) \qquad (i = 1, \cdots, m), \tag{7.1}$$

and the initial conditions $z_i(x_1) = z_{i1}$, the range of x being from x_1 to x_2. When a function $y = y(x)$ is given, the above equations determine uniquely m functions, $z_i(x)$. A curve of the class K which is to be considered, is represented by an equation

$$C: \quad y = y(x). \tag{7.2}$$

All curves of the class join two given points (x_1, y_1) and (x_2, y_2) in the xy-plane. A curve belongs to the class, if when substituted with its derivatives in equations (7.1), the resulting functions $z_i(x)$ satisfy not only the prescribed initial conditions at x_1, but also take on specified values $z_i(x_2) = z_{i2}$ at the other end point x_2.

A property of the curve C which will be found to be useful in the applications of Fischer's generalized derivative is that of normality. A curve C is said to be normal if there are no solutions λ_i of the equations

$$\lambda_i'(x) + \sum_1^m \frac{\partial g_j}{\partial z_i} \lambda_j(x) = 0,$$

which are also solutions of the equation

$$\sum_1^m \left\{ \lambda_j \frac{\partial g_j}{\partial y} - \frac{d}{dx}\left(\lambda_j \frac{\partial g_j}{\partial y}\right)\right\} = 0$$

except the solution $\lambda_i = 0$. The arguments $y(x)$, $y'(x)$, $z_i(x)$ of the derivatives of g in these equations are those belonging to C.

In order to define, after the manner of Volterra, a derivative of the functional $L(C)$ at a value $x = \xi$ and relative to the set of values $\xi_1, \cdots, \xi_m$ it is necessary to compare the value of the functional $L(C)$ on C with its value taken along a neighboring curve of the class K, defined by an equation

$$y = y(x) + \delta y(x),$$

where δy must be properly chosen in order that this curve will belong to the set K. Fischer [19, p. 379] has devised a method for constructing these functions $\delta y(x)$. Suppose that $\eta(x)$ is a continuous function with continuous first and second derivatives such that the conditions

$$|\eta(x)| \leq \delta, \qquad |\eta'(x)| \leq \delta, \qquad |\eta''(x)| \leq \delta,$$

(7.3)

$$\eta(x_1) = \eta(x_2) = 0,$$

are satisfied. A curve C_ε defined by the equation

$$C_\varepsilon: \qquad y = y(x) + \eta(x) + \sum_1^m \eta_i(x),$$

will belong to the class of curves K if the functions $\eta_i(x)$ are properly chosen. When the set of points ξ_i is arbitrarily chosen, one may determine the $\eta_i(x)$ as functions which have the same continuity properties as $\eta(X)$, which are identically zero everywhere except on intervals $\xi_i - \varepsilon \leqq x \leqq \xi_i + \varepsilon$ where they are of constant sign, and which are such that the curve C_ε is in K.

In defining the derivative of a function of a line mentioned in the last paragraph, the points ξ_i are considered as fixed. The variation $\eta(x)$ is chosen identically zero everywhere except on an interval $\xi - \varepsilon \leqq x \leqq \xi + \varepsilon$ where ξ is a fixed point, and also satisfies the hypotheses indicated above. If L(C) is a function of a line such that the limit

$$\lim_{\varepsilon \to 0} \frac{L(C_\varepsilon) - L(C)}{\sigma} = L'(C, \xi; \xi_1, \cdots, \xi_m),$$

exists uniquely, where σ is defined by the equation

$$\sigma = \int_{\xi-\varepsilon}^{\xi+\varepsilon} \eta(x)\, dx,$$

then that limit is called by Fischer the generalized Volterra derivative of the functional L(C) at the point ξ relative to the set of points $\xi_1, \cdots, \xi_m$ and the class of curves K. The derivative as defined above is said to be approached with order two because of the restriction $|\eta''(x)| < \delta$ imposed above.

In order to define a derivative which will be approached with order r, the inequalities (7.3) are replaced by

$$|\eta(x)| \leq \delta, \quad |\eta'(x)| \leq \delta, \quad \cdots, \quad |\eta^{(r)}(x)| \leq \delta, \tag{7.4}$$

and it is assumed that y(x) has continuous derivatives up to and including those of order r. A discussion similar to the one given on the preceding pages will still hold, and will again imply the existence of a derivative.

This extended derivative is defined for curves having at least continuous first and second derivatives, and not for all continuous curves. In order to get a more widely applicable derivative Fischer proved another theorem. In addition to the preceding hypotheses it is presupposed that there is a neighborhood of C such that if C'_ε, C' are in that neighborhood then the inequality

$$\left| \frac{L(C'_\varepsilon) - L(C')}{\sigma} - L'(C', \xi) \right| < \varepsilon$$

is always satisfied.

The comparison curves for this case have an equation of the form

$$C_{\varepsilon\alpha}: \qquad y = y(x) + \eta(x,\alpha) + \sum_1^m \eta_i(x,\alpha),$$

where the function $\eta(x,\alpha)$ is given in advance and the functions $\eta_1, \cdots, \eta_m$ are determined as before as functions of x and the parameter α. It is also assumed that the

functions $\eta, \eta', \dots, \eta^{(n)}$ approach zero uniformly with α and that the function $\eta(x,\alpha)$ has at $\alpha = 0$ a derivative $\eta_\alpha(x,0)$ with respect to α which is continuous in x and approached uniformly with respect to x. Under these conditions the functions $\eta_i(x,\alpha)$ may be determined with properties similar to those imposed in the preceding paragraphs and such that the curve $C_{\epsilon\alpha}$ belongs to the class K.

If these assumptions are fulfilled, then the limit

$$\lim_{\epsilon \doteq 0} \frac{L(C_{\epsilon\alpha}) - L(C)}{\alpha}$$

is unique, and can be proved to have the form

$$\lim_{\epsilon \doteq 0} \frac{L(C_{\epsilon\alpha}) - L(C)}{\alpha} = \int_a^b L'(C,x)\,\eta_\alpha(x,0)\,dx.$$

In particular, if $\eta(x,\alpha)$ is an admissible variation, that is if the curve

$$C_\alpha: \qquad y = y(x) + \eta(x,\alpha),$$

belongs to the set K, then $L(C_\alpha)$ is a uniquely determined function of α and its differential at $\alpha = 0$ has the form

$$\left[\frac{dL(C_\alpha)}{d\alpha}\right]_{\alpha=0} = \int_a^b L'(C,x)\,\eta_\alpha(x,0)\,dx. \tag{7.5}$$

This theorem on the generalized Volterra derivative is of a type applicable in the calculus of variations. More conditions than Volterra used are imposed on the curves under consideration, but the results obtained are more useful than those of Volterra in the applications.

8. <u>Fischer's derivative of a function of a surface.</u> Fischer has defined a derivative of a function of a surface in the Volterra manner, extending the ideas described in the last section to functions of surfaces in three-space. He also defined a derivative of a function of a surface relative to a special class of surfaces K. These theories will be considered in this section.

The derivative of a function of a surface L(S) is defined in an article in the American Journal of Mathematics [20, p. 289]. The surfaces to be considered are represented by equations of the form

$$S: \qquad z = z(x,y),$$

$$S_\epsilon: \qquad z = z(x,y) + \eta(x,y),$$

and are defined over a region R of the xy-plane. The functions $z(x,y)$ and $\eta(x,y)$ are continuous and have continuous derivatives up to and including those of order r. The function η vanishes for all values of x and y outside a region interior to R defined by the inequalities $x_0 - \epsilon < x < x_0 + \epsilon'$, $y_0 - \epsilon < y < y_0 + \epsilon$. In this region η has a constant sign. The absolute values of the partial derivatives of $\eta(x,y)$ up to and including those of order r are

supposed to be everywhere less than ε. The derivative of a function of a surface $L(S)$ at the point x_0, y_0 is then defined to be the limit

$$L'(S; x_0, y_0) = \lim_{\varepsilon \doteq 0} \frac{L(S_\varepsilon) - L(S)}{\sigma},$$

if the limit exists, where

$$\sigma = \int_{x_0-\varepsilon}^{x_0+\varepsilon} \int_{y_0-\varepsilon}^{y_0+\varepsilon} \eta(x,y)\, dy\, dx.$$

Such a derivative is said to be approached with order r because of the restrictions above placed upon η and its derivatives.

If it is also assumed that the derivative $L'(S, x_0, y_0)$ is continuous with order r in its arguments S, x_0, y_0 and is approached uniformly with order r for all (x_0, y_0) in R and surfaces S in a neighborhood of an initial one, then a formula may be deduced for the derivative with respect to α of the functional $L(S_\alpha)$ on a one-parameter family of surfaces S_α, α being the parameter of the family. The family of surfaces S_α may be defined by an equation

$$S_\alpha: \qquad z = z(x,y) + w(x,y,\alpha),$$

where $w(x,y,\alpha)$ has the same continuity properties as $z(x,y)$ and the derivatives of w with respect to x and y vanish on the boundary of R and approach zero uniformly in x and y when α approaches zero. It is also presupposed that

$w(x,y,\alpha)$ has a derivative with respect to α such that $w_\alpha(x,y,0)$ exists, is continuous, and approached uniformly for all (x,y) in R. If the functional $L(S)$ and the surfaces S_α satisfy the assumptions stated above, then it turns out that the derivative of $L(S_\alpha)$ with respect to the parameter α is given by the relation

$$\left[\frac{d\, L(S_\alpha)}{d\alpha}\right]_{\alpha=0} = \iint_R L'(S;x,y)\, w_\alpha(x,y,0)\, dx\, dy \tag{8.1}$$

This work of Fischer's on the derivative of a function of a surface is an extension to functions of surfaces in three-space of Volterra's ideas on derivatives of functions of lines, as indicated in the first two sections of this chapter. The applications of this derivative to problems of minimizing functionals will be considered in Section 18.

Fischer [20, p. 297] also defined a derivative of a function of a surface relative to a class of surfaces K. In doing so he considered m functionals $M_i(S)$ which are continuous and whose derivatives $M_i'(S,x,y)$ are continuous and approached uniformly with order r in a neighborhood of the surface S. It is also presupposed that there are some points such that the inequality

$$\left| \begin{matrix} M_j'(S;x_i,y_i) \\ (i,j=1,\cdots,m) \end{matrix} \right| \neq 0$$

is satisfied. A surface S^1 is said to belong to the set K

if it satisfies the relation

$$M_j(S^1) - M_j(S) = 0 \qquad (j = 1, 2, \cdots, m).$$

If a surface belongs to the set K, there is no assurance that a neighboring surface S_0

$$S_0: \quad z = z(x, y) + \eta(x, y),$$

will belong to the set. However, Fischer shows in a manner similar to that used for the plane case (see page 25) that if m points (x_i, y_i) are properly chosen, then m functions $\eta_i(x, y)$ having properties similar to those of $\eta(x, y)$ may be determined such that the surface

$$S_m: \quad z = z(x, y) + \eta(x, y) + \sum_1^m \eta_i(x, y),$$

belongs to the set K.

For this restricted class of surfaces Fischer defines his derivative as the limit

$$L'(S; x_0, y_0, x_i, y_i) = \lim_{\epsilon \doteq 0} \frac{L(S_m) - L(S)}{\sigma},$$

where σ has the value indicated on page 29. This limit is called the derivative of the functional L(S) at the point (x_0, y_0) relative to the set of surfaces K and the points (x_i, y_i). The functions M_i appear in the value of the derivative L'(S). For example, L' may be written in the form

(8.2)
$$L'(S;x_0,y_0,x_i,y_i) = \lim_{\varepsilon \doteq 0}\left[\frac{L(S_0)-L(S)}{\sigma} + \sum_1^m \frac{L(S_i)-L(S_{i-1})}{\sigma_i}\frac{\sigma_i}{\sigma}\right]$$
$$= L'(S;x_0,y_0) + \sum_1^m L'(S;x_i,y_i)\lim_{\varepsilon = 0}\frac{\sigma_i}{\sigma},$$

where σ_i is defined by the relation

$$\sigma_i = \int_{x_i-\varepsilon}^{x_i+\varepsilon}\int_{y_i-\varepsilon}^{y_i+\varepsilon}\eta_i(x,y)\,dy\,dx.$$

The limit σ_i/σ may be evaluated in terms of the functions M_i. Since S_m belongs to the set K, it follows that the equations

$$\frac{M_j(S_m) - M_j(S)}{\sigma} = 0,$$

are satisfied, and hence as ε and σ approach zero it is evident that

(8.3) $$M_j(S;x_0,y_0) + \sum_1^m M_j(S;x_i,y_i)\lim_{\varepsilon \doteq 0}\frac{\sigma_i}{\sigma} = 0.$$

If these equations are solved for the desired ratios and the results substituted in equation (8.2), it can be shown [20, p. 301] that the derivative $L'(S)$ may be written in the form

(8.4) $$L'(S;x_0,y_0,x_i,y_i) = L'(S;x_0,y_0) + \sum_1^m \lambda_K M'_K(S;x_0,y_0),$$

where the λ_K depend only upon the derivatives of L and M_i on the surface S at the points (x_k,y_k).

These results are a generalization of those obtained

by Fischer for the analogous plane case, and are especially applicable to the problem of minimizing a functional whose arguments are subjected to certain restrictions. The applications of such derivatives to the problem of Lagrange in the calculus of variations will be discussed later.

In another article Fischer [31, p. 259] considered surfaces with exceptional points or curves. The type of hypotheses used in the preceding papers on the surfaces

$$
\begin{aligned}
S: &\quad z = z(x,y),\\
S_\varepsilon: &\quad z = z(x,y) + \eta(x,y),\\
S_\alpha: &\quad z = z(x,y) + w(x,y,\alpha),
\end{aligned}
$$

and the functionals $L(S)$ are adopted, except that there is an exceptional point (x_0, y_0) or an exceptional curve. Fischer's notations and arguments are lacking in clearness. The following paragraphs give the hypotheses and results of his paper with modifications which seem necessary in order to harmonize them with the text of his proofs.

The first type of exceptional point (x_0, y_0) which he considers is one at which the limit

$$
\lim_{\varepsilon \doteq 0} \frac{L(S_\varepsilon) - L(S)}{\alpha} = 0, \tag{8.5}
$$

holds for every $\eta(x,y)$ dependent upon α, and such that the ratio $\eta(x,y)/\alpha$ is bounded for all points (x,y) in R and all values of α near $\alpha = 0$, and furthermore such that the limit

$$\lim_{\varepsilon \doteq 0} \frac{1}{\alpha} \frac{\partial^{i+j} \eta(x_0, y_0)}{\partial x^i \partial y^j},$$

exists. Fischer does not describe explicitly the dependence of η (x,y) upon α, or of α upon ε but it appears in his proofs that η is always a function $\eta(x, y, \alpha)$ whose absolute value approaches zero uniformly with α for all points (x,y) in R. The inequality $|\eta(x,y)| < \varepsilon$ can be attained by making α approach zero. His limits as ε approaches zero should perhaps therefore be interpreted to be limits as α approaches zero. For an exceptional point (x_0, y_0) of the type described above, it is shown that the form of the derivative as given by the expression (8.1) is unchanged.

A second type of exceptional point (x_0, y_0) is one for which there exists a set of constants a_{ij} such that the relation

$$\lim_{\varepsilon = 0} \frac{L(S_\varepsilon) - L(S)}{\alpha} = \sum_{i,j=0}^{m} a_{ij} \lim_{\varepsilon \doteq 0} \frac{1}{\alpha} \frac{\partial^{i+j} \eta(x_0, y_0)}{\partial x^i \partial y^j}, \tag{8.6}$$

holds. For such points, by a method similar to that used by Volterra (see page 10) Fischer shows that the derivative of $L(S_\alpha)$ with respect to α for $\alpha = 0$ may be expressed in the form

$$\left[\frac{dL(S_\alpha)}{d\alpha}\right]_{\alpha=0} = \iint_R L'(S; x, y)\, w_\alpha(x, y, 0)\, dx\, dy + \sum_{i,j=0}^{m} a_{ij} \frac{\partial^{i+j} w_\alpha(x_0, y_0, 0)}{\partial x^i \partial y^j}. \tag{8.7}$$

In order to prove the above formula it is assumed that the relation (8.6) holds at the point (x_0, y_0). A new function $\bar{L}(S)$ is defined by the expression

$$\bar{L}(S) = L(S) - \sum_{i,j=0}^{n} a_{ij} \frac{\partial^{i+j} z(x_0, y_0)}{\partial x^i \partial y^j}.$$

This function $\bar{L}(S)$ will satisfy equation (8.5) and have a derivative in the form given by the expression (8.1). The derivatives $\bar{L}'(S;x,y)$ and $L(S;x,y)$ are equal except at the exceptional points. The combination of these results gives the derivative (8.7) for $L(S)$.

Fischer obtains a second derivative for functionals of a surface with an exceptional point (x_0, y_0) of the second kind described above in the form

$$\left[\frac{d^2 L(S_\alpha)}{d\alpha^2}\right]_{\alpha=0} = \iint_R \Big\{ \iint_R L''(S; x_0, y_0, x, y) w_\alpha(x_0, y_0, 0)\, w_\alpha(x, y, 0)\, dx\, dy$$

(8.8)
$$+ \sum_{i,j=0}^{n} a_{ij}(x_0, y_0) \frac{\partial^{i+j} w_\alpha(x_0, y_0, 0)}{\partial x^i \partial y^j} w_\alpha(x_0, y_0, 0)$$
$$+ L'(S; x_0, y_0)\, w_{\alpha\alpha}(x_0, y_0, 0)\Big\}\, dx_0\, dy_0,$$

the a_{ij} being constants whenever the point (x_0, y_0) is fixed.

For functions of surfaces with exceptional curves Fischer [31, p. 262] obtains similar forms for the derivative of a functional $L(S)$. Two cases are considered, the first being for those curves along which a certain limit

$$\lim_{\varepsilon \doteq 0} \frac{L(S_\varepsilon) - L(S)}{\varepsilon \alpha} = 0,$$

holds uniformly in a neighborhood of the given surface. For such curves the form of the derivative as given by the expression (8.1) is unchanged.

The second type of exceptional curve is one for which there exists a set of n continuous functions $a_j(s)$ such that along the curve

$$\lim_{\varepsilon \doteq 0} \frac{L(S_\varepsilon) - L(S)}{\varepsilon \alpha} = \sum_{j=0}^{n} \frac{1}{\varepsilon\alpha} \int_{s-\varepsilon}^{s+\varepsilon} a_j(s) \frac{\partial^j \eta(s,0)}{\partial n^j}\, ds,$$

where the variation η (s,n) is expressed in terms of the arc length s and the normal distance to the curve. By a proof similar to the analogous case for surfaces which have exceptional points of the second kind Fischer shows that

$$\left[\frac{d\,L(S_\alpha)}{d\alpha}\right]_{\alpha=0} = \iint_R L'(S;x,y)\, w_\alpha(x,y,0)\, dx\, dy \tag{8.9}$$
$$+ \sum_{j=0}^{n} \int_a^b a_j(s) \frac{\partial^j w_\alpha}{\partial n^j}\, ds.$$

If the functional L(S) depends on a surface represented in parametric form by the equations

$$x = x(u,v), \qquad y = y(u,v), \qquad z = z(u,v),$$

Fischer [32, p. 123] defined the derivative of the functional to be

$$L'(S;u,v) = \lim_{\varepsilon \doteq 0} \frac{L(S_\varepsilon) - L(S)}{\sigma},$$

where

$$\sigma = \int_{u_0-\varepsilon}^{u_0+\varepsilon} \int_{v_0-\varepsilon}^{v_0+\varepsilon} \eta(u,v)\, H(u,v)\, du\, dv,$$

and where H du dv is the element of surface area. The customary hypotheses on the surfaces and functionals under consideration are understood to be fulfilled.

Fischer moulded his definitions and theorems so that they might be applicable to problems of the calculus of variations, and used them to deduce some necessary conditions for a minimum of a functional in general. These applications will be considered in a later section.

9. Definitions of the derivative of a functional by Pascal and Daniele. Other definitions of the derivative of a functional in the manner of Volterra have been made by Pascal and Daniele. Daniele [24, p. 319; 25, p. 496] considered the derivatives and differentials of a functional F whose argument y(x) depended on another function $\psi(\xi)$, and a value x of the independent variable. He shows in the manner of Volterra that if

$$F = F\left|\left[y\left(\overset{t}{\underset{a}{x}}\right) \right]\right|, \qquad y = y\left|\left[\psi\left(\overset{1}{\underset{0}{\xi}}\right), x \right]\right|,$$

the derivative of F is expressible in the form

$$F'\left|\left[\psi(\xi), x\right]\right| = \int_a^t F'\left|\left[y(x, z)\right]\right| y'\left|\left[\psi(\xi), z, x\right]\right| dz.$$

The functions under consideration are assumed to be continuous, and the derivatives F' and y' are taken in the Volterra sense. Analogous formulae are deduced when either F involves higher derivatives of y, or y involves those of ψ. The whole procedure is akin to the ordinary process

of differentiation of a function of a function.

Pascal [21, p. 40; 22, p. 68; 29, p. 318] defined the various derivatives of a functional in a manner slightly different from that of Volterra. The functionals he considers depend on a curve $y = y(x)$, and a variable t,

$$F = F([y], t), \tag{9.1}$$

where the range of x is $a \leq x \leq b$, and the range of t is $\alpha \leq t \leq \beta$.

Pascal distinguishes two types of functions of lines, the first of which is a functional, (9.1), independent of t which has a well determined value for all curves in a certain region of the xy-plane. Such a functional is called a point function of a line. The second type of functional, depending upon t as well as $y(x)$, is called a line function of a line. The functional F given by equation (9.1) is a point function of a line when t is kept fixed, and is a line function of a line when its value for all values of t are under consideration.

In order to determine the derivative of a point function of a line, Pascal first fixes the variable t in equation (9.1). His method of procedure is like that which Volterra used, (see page 6). On varying the function $y(x)$ in a neighborhood of a fixed abscissa x_1, a change ΔF of F is obtained. If the area between the original curve $y = y(x)$ and the corresponding altered curve is denoted by σ, the limit of the ratio $\Delta F/\sigma$ as the area σ ap-

proaches zero in the manner specified by Volterra, is called the partial derivative of the point function of a line with respect to the point x_1, and is denoted by $F'_{y(x_1)}$ or $\partial F/\partial y(x_1)$. This is Volterra's derivative with a different notation.

The partial derivative of a line function of a line with respect to a curve $y = y(x)$ and a particular ordinate $y(x_1)$ is denoted by the symbol $\partial[F]/\partial[y(x_1)]$. It is defined to be the limit of the ratio of the increment of area of the curve defined by F in the tF-plane, to the increment σ of area, near the ordinate x_1 of the curve defined by $y(x)$ in the xy-plane, as σ approaches zero in the way described in the definition of Volterra's derivative. The value of this derivative is given by the formula

$$\frac{\partial[F]}{\partial[y(x_1)]} = \int_\alpha^\beta F'_{y(x_1)}([y], t)\,dt, \tag{9.2}$$

where $F'_{y(x_1)}$ is the partial derivative of the function $F([y], t)$ thought of as a point function of a line as defined above for a fixed t. This Pascal derivative is independent of the type of variation, subject to Volterra's restrictions, chosen for y. Daniele has commented upon this definition of a derivative, and his comments will be described in following paragraphs.

Pascal also deduces a formula for the total derivative of a line function of a line $F([y(x)], t)$, with respect to $y(x)$. Suppose the function $y(x)$ is given an increment Δy

for each x. Then the functional F is changed into $F + \Delta F$. The derivative of F with respect to y, if it exists, is defined to be

$$\lim \frac{\int_{\alpha}^{\beta} \Delta F \, dt}{\int_{a}^{b} \Delta y \, dx},$$

as the denominator approaches zero, with the understanding that the mode of approach of Δy to zero is specified in advance. Pascal makes this specification by assigning $\Delta y = y(x, \varepsilon) - y(x)$, where $y(x,0) = y(x)$, and takes the limit as ε tends towards zero. In contrast with Pascal's definition of the partial derivative of a point function of a line, the variation of y(x) extends over the entire range of x, not being restricted to the neighborhood of a fixed point, and the value of the derivative depends in general upon the special mode chosen for Δy to approach zero.

The value of the limit so defined, if it exists, is designated by $d[F]/d[y]$. Pascal shows that it may be represented in the form

$$(9.3) \quad \frac{d[F]}{d[y]} = \lim_{\varepsilon \doteq 0} \frac{\int_{\alpha}^{\beta} \frac{\Delta F}{\varepsilon} dt}{\int_{a}^{b} \frac{\Delta y}{\varepsilon} dx} = \frac{1}{\int_{a}^{b} \frac{\partial y}{\partial \varepsilon} dx} \int_{\alpha}^{\beta} \int_{a}^{b} F'_{y(x)}\left([y(x)] t\right) \frac{\partial y}{\partial \varepsilon} dx \, dt.$$

In securing this result, he uses Volterra's formula for the limit of the ratio $\Delta F/\varepsilon$ as ε approaches zero. It should be noted that the formula (9.3) depends upon the mode of representation of y as a function of the parameter ε.

Daniele [27, p. 102] has shown that if the func-

tional F(y,t) does not contain the variable t, then formulae (9.2) and (9.3) reduce to

$$(\beta-\alpha)F'_{y(x_1)}, \qquad (\beta-\alpha)\frac{\int_a^b F'_{y(x)}\frac{\partial y}{\partial \varepsilon}dx}{\int_a^b \frac{\partial y}{\partial \varepsilon}dx},$$

respectively. Thus, the partial and total derivatives of Pascal for this special case are equal to the corresponding derivatives of Volterra multiplied by a constant.

For the total derivative, Daniele [27, p. 106] showed that as y(x) does not depend upon t, formula (9.3) may be expressed in the form

$$\frac{d[F]}{d[y]} = \frac{\int_a^b \frac{\partial y}{\partial \varepsilon}\int_\alpha^\beta F'_{y(x)}\,dt\,dx}{\int_a^b \frac{\partial y}{\partial \varepsilon}dx} = \frac{\int_a^b \frac{\partial [F]}{\partial [y(x)]}\frac{\partial y}{\partial \varepsilon}dx}{\int_a^b \frac{\partial y}{\partial \varepsilon}dx}.$$

If it is presupposed that $\partial y/\partial\varepsilon$ has a fixed sign for $a \leqq x \leqq b$, it follows that

$$\frac{d[F]}{d[y]} = \left(\frac{\partial[F]}{\partial[y(x)]}\right)_{x=x_0},$$

where x_0 is a suitably selected mean value on the interval ab. This formula can be considered as a mean value relation for the Pascal derivatives. It shows that Pascal's total derivative is always equal to the value of Pascal's partial derivative of a line function of a line at a suitably chosen value x_0 on the interval ab. For the special case when F([y] ,t) does not contain t explicitly, the last formula and (9.2) show that

$$\frac{d[F]}{d[y]} = (\beta - \alpha) F'_{y(x_0)}([y]),$$

so that Pascal's total derivative is in this case equal to a mean value of Volterra's derivative of F multiplied by the constant $(\beta - \alpha)$.

10. Miscellaneous papers concerning Volterra's derivative of a functional. The topics discussed in this section are the work of Daniele on solving certain functional equations, Mandelbrojt's study of a type of functional derivative, and the results of a number of articles on the expansions of functionals in series, analogous to Taylor's series.

Daniele [26] considered functionals of the type

$$f(x) = f\Big|\Big[\varphi(\overset{x}{\underset{0}{\xi}}), x\Big]\Big|, \tag{10.1}$$

where f is defined as a function of φ on the interval $0 \leq \xi \leq x$, and of the variable x. He also thought of this equation as defining φ as a function $\varphi\big|[f(\xi), x]\big|$ and studied not only the relations between f and φ , but also those existing between the derivatives of f with respect to φ and φ with respect to f. He developed this analysis first for the particular functionals

$$\begin{aligned} f(x) &= \varphi(x) + \int_0^x \Phi(x,y)\,\varphi(y)\,dy, \\ f(x) &= \int_0^x \Phi(x,y)\,\varphi(y)\,dy, \end{aligned} \tag{10.2}$$

and later considered briefly its extension for functionals of more general types. For functionals given by the first of equations (10.2), it follows that φ may be expressed

in the form

$$\varphi(x) = f(x) + \int_0^x F(x,z)\, f(z)\, dz,$$

where F(x,z) is the resolvent kernel of Φ (x,y). Hence one finds that

$$f'_\varphi = f'|[\varphi(y), x, \eta]| = \Phi(x,\eta),$$

$$\varphi'_f = \varphi'|[f(y), x, \eta]| = F(x,\eta),$$

Daniele then states the principle of reciprocity

$$\Phi(x,t) + F(x,t) = -\int_t^x \Phi(x,z)\, F(z,t)\, dz$$

$$= -\int_t^x F(x,z)\, \Phi(z,t)\, dz,$$

well known in integral equation theory, with the aid of which he then shows that

$$f'_\varphi + \varphi'_f = -\int_t^x f'_\varphi(x,z)\, \varphi'_f(z,t)\, dz.$$

An immediate consequence of this relation is that the derivatives f'_φ and φ'_f are permutable in the sense of Volterra's first type of permutability [17, p. 133].

Daniele then shows that such ideas of reciprocity and permutability can be extended to more general functionals of the type (10.1), in particular to those which have a variation in the form

$$\delta f(x) = \int_0^x f'\big|[\varphi(\xi), x, \eta]\big|\,\delta\varphi(\eta)\,d\eta + a(x)\,\delta\varphi(x),$$

where a(x) is a function independent of φ .

Mandelbrojt [37, p. 151] defined a derivative of order α of a functional f(x), α being any real number, by means of the relations

$$D_x^0 f(x) = f(x),$$

$$D_x^\alpha f(x) = \frac{1}{\Gamma(-\alpha)} \int_{+1}^x (x-z)^{-\alpha-1} f(z)\,dz \qquad (\alpha < 0),$$

$$D_x^\alpha f(x) = \frac{d^p}{dx^p} D_x^{\alpha-p} f(x) \qquad (0 < p-1 < \alpha \leq p).$$

Let us for convenience denote the functional derivative D_x^α f(x), a function of the two variables x and α , by f(x, α). With the help of this notation, an integral

$$(10.3) \quad I = \int_0^1 F\left[x, f(x), f'(x), f''(x), \cdots, f^{(n)}(x)\right] dx,$$

of the calculus of variations may be written in the form

$$(10.4) \quad I = \int_0^1 F\big|[x, f(x,\alpha)]\big|\,dx, \qquad (0 \leq \alpha \leq T).$$

Mandelbrojt then shows that the variation δI of the functional (10.4) may be represented by the equation

$$\delta I = \int_0^1 \left\{ \delta f(x) \int_0^T (-1)^\xi D_x^\xi F'_\alpha \big|[x, f(x,\alpha), \xi]\big|\,d\xi \right\} dx,$$

where the integrand F_α' is the Volterra derivative of F with respect to f at the point ξ for the value α. An analogue of the Euler-Lagrange equation of the calculus of variations for the problem of minimizing the integral (10.3) is then obtained for the problem of minimizing the functional (10.4) by setting

$$F_\alpha' \left| \left[x, f(x,\alpha), \xi \right] \right| = 0$$

A study of the properties and solutions of equations of this type has not been made as yet, but Mandelbrojt states that he will discuss this question in a later note.

Various generalizations of Taylor's series have been made with the help of the idea of the Volterra derivative. Volterra gives for a functional $F[y]$ having derivatives up to and including those of order n the expansion [17, p. 28; 38, p. 26]

$$\begin{aligned} F[y+\eta] &= F[y] \\ &+ \sum_1^{n-1} \frac{1}{i!} \int_a^b \cdots \int_a^b F^{(i)}[y; \xi_1, \cdots, \xi_i] \eta_1(\xi_1) \cdots \eta_i(\xi_i) \, d\xi_1 \cdots d\xi_i \\ &+ \frac{1}{n!} \int_a^b \cdots \int_a^b F^{(n)}[y+\theta\eta; \xi_1, \cdots, \xi_n] \eta_1(\xi_1) \cdots \eta_n(\xi_n) d\xi_1 \cdots d\xi_n, \end{aligned}$$

where $0 \leq \theta \leq 1$, and the $F^{(i)}$ are successive Volterra derivatives. This formula is analogous to Taylor's series with a remainder term. Similar expansions for analytic functionals have been made by Kakeya [33, p. 341; 34, p. 177] and Levy [36, p. 102].

CHAPTER II

THE FRÉCHET DIFFERENTIAL OF A FUNCTIONAL

Introduction. In 1909 F. Riesz showed that a linear functional with suitable continuity properties could always be represented as a Stieltjes integral. About the same time a new conception of the differential of a functional was introduced by Fréchet. According to his definition, the differential was a linear functional, and could therefore, by the Riesz theorem, be represented as a Stieltjes integral.

In this chapter the theory of the representation of linear and bilinear functionals as Stieltjes integrals is considered as a basis for the introduction of the Fréchet type of differential. The work of Fréchet and Gateaux on the definitions of the differentials of a functional, together with the modification of these definitions by Le Stourgeon in her study of the problem of minimizing a functional, are the topics treated in this chapter.

11. The Riesz theorem. F. Riesz [39, p. 974] in the Comptes Rendus first showed that a linear functional F(y), having continuity of order zero, could always be represented as a Stieltjes integral in the form

$$F(y) = \int_a^b y(x)\, du(x). \tag{11.1}$$

A more detailed proof was published in the Annales Scientifiques de L'École Normale Superieure [41, p. 33]. Later Riesz [47, p. 2] proved the same theorem under slightly different hypotheses, using a modularity property instead of

the continuity of the linear functional previously required.

Bliss, in his unpublished lecture notes of courses on functions of lines in 1916 and 1920 at the University of Chicago, proved this theorem by methods similar to those which Riesz used. L. M. Graves, in his course on the same subject in 1929 generalized the integral used to that of Lebesgue-Stieltjes; extended the class of functions y, for which F is defined to the Borel measurable class; and proved certain relations existing between such generalized representations.

One of the most concise and satisfactory published proofs of the theorem of Riesz was made by Hahn [55, p. 535]. Since the functional Hahn used is defined only for continuous functions y, his theorem is not as general as those which Riesz and Graves proved, but it is ample for most of the applications one desires to make in the calculus of variations.

Another proof of the theorem is given in Levy's text [36, p. 55], but follows closely the method originally used by Riesz. Hildebrandt [51, p. 183], in an article on integrals related to those of Lebesgue, states the Riesz theorem, and gives some of the consequences of it as seen from the Stieltjes integral standpoint.

12. Extensions of the Riesz theorem. The Riesz theorem, that a linear functional may be represented as a Stieltjes integral, has been extended in two directions, one of which is for functionals F which depend on several functions $y_i(x)$, the other being for the case when y is a func-

tion of n independent variables x_i.

For this latter case, Fischer [49, p. 640; 50, p. 37] and Fréchet [40, p. 241] have shown that a linear functional F [$y(x_1, \cdots, x_n)$] , where y is a continuous function, may be represented in the form

$$F\left|\left[y(x_1, \cdots x_n)\right]\right| = \int_T f(x_1, \cdots, x_n)\, d\,\alpha(x_1, \cdots, x_n), \tag{12.1}$$

where T is the region in which the functions y(x) are defined and where α is a function of limited variation, suitably defined in n-dimensional space, the integral being taken in the sense of Stieltjes. This is a generalization of the Riesz theorem to n dimensions.

For a linear functional F [$y_i(x)$] which depends on n independent functions $y_i(x)$, Hahn [55, p. 537] has shown that F [$y_i(x)$] may be represented as a sum of Stieltjes integrals

$$F\left[y_i(x)\right] = \sum_{i=1}^{n} \int_a^b y_i(\xi)\, d\alpha_i(\xi) \tag{12.2}$$

The sum of integrals appears because the linear functional F is expressible in the form

$$F\left[y_1, \cdots, y_n\right] = F\left[y_1, 0, 0, \cdots\right] + F\left[0, y_2, 0, 0, \cdots\right] + \cdots + F\left[0, 0, \cdots, 0, y_n\right].$$

The α_i which occur in the expression (12.2) are of limited variation. Hahn also proves that under suitable conditions $F\left[y_i(x)\right]$ can be expressed as the sum

$$F\left[y_i(x)\right] = \sum_{i=1}^{n} \int_a^b y_i(\xi)\,\beta_i(\xi)\,d\xi,$$

where the functions β_i are integrable in the sense of Lebesgue.

The representation of a bilinear functional $B\left[f,g\right]$, linear in each of the variables f and g separately and with suitable continuity properties, was considered by Fréchet [48, p. 215]. If the functions f and g are continuous and well defined over the region $a \leq x \leq a'$, $b \leq y \leq b'$, then the functional is expressible in any one of the forms

$$B[f,g] = \int_a^{a'} f(x)\,d_x \int_b^{b'} g(y)\,d_y\,\alpha(x,y),$$

$$B[f,g] = \int_b^{b'} g(y)\,d_y \int_a^{a'} f(x)\,d_x\,\beta(x,y),$$

$$B[f,g] = \int_a^{a'} \int_b^{b'} f(x)\,g(y)\,d_x\,d_y\,\gamma(x,y).$$

The functions α, β, γ which Fréchet used are defined over the same region as f and g, are not zero for $x = a'$, or $y = b'$, and are regular and of bounded variation with respect to one or both of their arguments. The idea of a bilinear functional occurs in the study of the second differential of a functional, as we shall see presently.

Wong [57, p. 34] has extended the Stieltjes integral representation of bilinear functionals to the case when the functions f and g depend upon more than one independent variable. He obtained for bilinear functionals results similar to those deduced by Fischer and Fréchet for linear functionals. Thus, for a functional B [f(x), g(y)] where $x = (x_1, \cdots, x_n)$, $y = y(y_1, \cdots, y_n)$ Wong shows that B may be represented in any one of the forms

$$B[f,g] = \int_T f(x)\, d_{(m)} \int_S g(y)\, d_{(n)}\, \alpha(x,y)$$

$$= \int_R f(x)\, g(y)\, d_{(n+m)}\, \alpha(x,y),$$

$$B[f,g] = \int_S g(y)\, d_{(n)} \int_T f(x)\, d_{(m)}\, \beta(x,y)$$

$$= \int_R g(y)\, f(x)\, d_{(n+m)}\, \beta(x,y),$$

where R, S, T are suitably chosen regions, α and β satisfy certain restrictions regarding their variation, and the integrals are taken in the sense of Stieltjes.

The theory of the representation of linear and bilinear functionals as Stieltjes integrals will be useful in the application of the theory of functionals to the calculus of variations.

13. The Fréchet differentials. In this section the notions of Fréchet on differentials of functions and functionals will be considered. In his first article [42, p. 385] Fréchet points out that the differential, at a parti-

cular point, of a function of several variables is a linear function of the increments of the arguments of the function, and that it differs from the increment of the original function by a quantity which approaches zero with the modulus of the increments of the arguments. In this article he considers only the case of a function of a real variable, but states that similar relations exist for functionals.

In a later article [44, p. 135] Fréchet applies this idea to functionals. The functionals $F(y)$ which he considers are defined over a field of continuous functions $y(x)$ on the interval ab. are continuous in this field, and such that $F(y_n(x))$ approaches $F(y)$ as $y_n(x)$ converges uniformly to $y(x)$. If a functional so defined is to have a differential at the argument y, then according to the definition of Fréchet, there will exist a linear functional $L(\eta)$, such that the relation

$$F(y+\eta) = F(y) + L(\eta) + \epsilon M(\eta),$$

is satisfied; where ϵ tends towards zero with the maximum of the absolute value of η, $M(\eta)$. The linear functional L so defined is said to be the differential of the function $F(y)$. Similarly, if the functional F depends on several arguments $y, z, \cdots, w$ and has properties analogous to those described above for the case of a single one, then the linear functional L will have as its arguments the increments of all the variables, and again may be defined to be the differential of the functional.

It is interesting to compare the differential of Fréchet with that of Volterra (see Sections 1, 2). Volterra defines first his derivative of a functional at a point, and then proceeds to derive an expression for the differential of the functional by analogy with the case of a function of a real variable. Fréchet, on the other hand, defines his differential directly without using the definition of a derivative. Both of these two authors presuppose continuity of order zero for the functional F, which is more than is possessed by the functionals of the calculus of variations.

After defining the first differential of a functional, Fréchet considers the problem of representing it as a Stieltjes integral by means of the Riesz theorem [44, p. 145]. As his differential is a linear functional, it may be so expressed. Several different representations are given, depending on the nature of the functional L and the type of discontinuities of the function of limited variation occurring in the differential.

Fréchet defines a second differential of a functional in the form of a bilinear functional [48, p. 232]. It is assumed that $F(y)$ has a first differential $L(y; \eta)$ corresponding to an increment η of $y(x)$. Fréchet further presupposes that $L(y; \eta)$ has a differential B with respect to the argument y for a new increment η_1. This new functional $B(y; \eta, \eta_1)$ is a bilinear functional in η and η_1, and is called the second differential of F. It might have

been defined directly from the original functional F by means of the relation

$$B(y;\eta,\eta_1) = \lim_{\lambda \to 0} \frac{F(y+\lambda(\eta+\eta_1)) - F(y+\lambda\eta) - F(y+\lambda\eta_1) + F(y)}{\lambda^2}$$

an expression which is obviously symmetric in its two arguments η, η_1. A bilinear functional may be represented as an iterated Stieltjes integral (see page 49). It is readily seen that higher differentials of a functional may be defined in a similar manner. In a later paper [56, p. 293], Fréchet considers definitions of differentials of functionals in general analysis.

14. Le Stourgeon's modification of the Fréchet differential. A definition of a differential of a functional, following the ideas of Fréchet, but designed especially to be applicable to the problems of the calculus of variations was made by Le Stourgeon [53, p. 357]. As continuity of an order higher than zero is required for such problems, the definition of a differential of a functional must necessarily be newly phrased.

Le Stourgeon considered an arc $y_0 = y_c(x)$ which is continuous and has continuous first derivatives on the interval ab, and a functional F(y) which is defined for functions y in a neighborhood of order one of y_0. The variation η of y is assumed to have the same continuity properties as y. The functional F(y) is then said to have a differential at y_0 if there exists a linear functional $L(\eta)$ with con-

tinuity of order one (see page 6) such that for all arcs $y_0 + \eta$ in a neighborhood of y_0 the relation

$$F(y_0+\eta) = F(y_0) + L(\eta) + M(\eta)\,\varepsilon(\eta),$$

holds where $M(\eta)$ is the maximum absolute value of η and η' on the interval ab and $\varepsilon(\eta)$ is a quantity vanishing with $M(\eta)$. Under these conditions the linear functional L is defined to be the differential of F, and may be represented by the Riesz theorem as the sum of two Stieltjes integrals

$$L(\eta) = \int_a^b \eta(x)\,d\alpha(x) + \int_a^b \eta'(x)\,d\beta(x),$$

where the functions $\alpha(x)$, $\beta(x)$ are of limited variation on the interval ab.

Le Stourgeon defined second differentials of a functional in a similar manner. The functional $F(y)$ is said to have a second differential at y_0 if there exists a linear functional L, and a bilinear functional $B(\eta, \eta_1)$ having continuity of order one on the interval ab and such that

$$F(y_0+\eta) = F(y_0) + L(\eta) + B(\eta,\eta) + M^2(\eta)\,\varepsilon(\eta),$$

for every arc $y_0 + \eta$ of the type previously described. As before, $M(\eta)$ represents the maximum absolute value of η and η' on the interval ab, and $\varepsilon(\eta)$ is a quantity which vanishes with $M(\eta)$. The value of the bilinear functional $B(\eta, \eta_1)$ when $\eta = \eta_1$ is called the second

differential of the functional F. The functional $B(\eta, \eta_1)$ may be expressed in the form

$$B(\eta,\eta_1) = \int_a^b \int_c^d \Big\{ \eta(x)\,\eta_1(y)\,d_{xy}\,p(x,y) + \eta'(x)\,\eta_1(y)\,d_{xy}\,q'(x,y) + \eta(x)\,\eta_1'(y)\,d_{xy}\,q''(x,y) + \eta'(x)\,\eta_1'(y)\,d_{xy}\,r(x,y) \Big\},$$

where the functions p,q',q",r are of limited variation with respect to the two variables of integration, x and y.

The work of Le Stourgeon stands in the same relation to the work of Fréchet as did Fischer's to Volterra's work. The purpose of both Le Stourgeon and Fischer was to extend definitions of differentials so that they would be applicable to the problems of the calculus of variations. Both authors deduced from their definitions of differentials necessary conditions for the existence of a minimum of a functional, which are of a type applicable to the problems of the calculus of variations.

15. Other definitions of differentials similar to those of Fréchet. The work of Gateaux and Levy on functionals, together with a definition of a differential of a functional by Bliss will be considered in this section.

Gateaux studied the problem of representing a functional as the limit of a sequence of other functionals each of which is expressible as a sum of integrals [43, p. 646; 45, p. 310; 46, p. 405; 54, p. 9]. These sums are analogous to polynomials in a finite number of variables, so that Gateaux's theorem is an analogue of the theorem of Weier-

strass concerning the approximation of a function of $x_1, \cdots, x_n$ by a polynomial in these variables. Gateaux obtained expressions for functionals $F\left|\left[y(x)\right]\right|$ with continuity of order zero in the form

$$F\left|\left[y(x)\right]\right| = \lim_{n\to\infty}\left\{K_{n_0} + \sum_{1}^{\infty}\int_{-n}^{n}\cdots\int_{-n}^{n}K_{n_p}(x_1,\cdots,x_p)\,\zeta_n(x_1)\cdots\zeta_n(x_n)\,dx_1\cdots dx_n,\right.$$

where

$$\begin{aligned}\zeta_n(x) &= y(x) \quad \text{if } |y(x)| \leq n,\\ &= n \quad \text{if } y(x) > n,\\ &= -n \quad \text{if } y(x) < -n,\end{aligned}$$

and the functions K_{ns} are properly chosen continuous functions. In another paper [52, p. 72] he deduced formulas by means of which continuous functions of a denumerable infinity of variables $x_1, x_2, \cdots$ could be approximated by polynomials in n of these variables, obtaining better approximations as n approached infinity.

The problem of determining the variation of a functional was also considered by Gateaux [52, p. 82; 54, p. 11] Levy, after Gateaux's death [36, p. 51] formulated another definition of the variation, or differential of a functional, following Gateaux's mode of attack. If, for example, $F(y)$ is a real continuous functional of a real continuous function $y(x)$, $0 \leq x \leq 1$, then the first variation of $F(y)$ is defined by Gateaux to be

$$\delta F(y,\eta) = \left[\frac{d}{d\lambda} F(y+\lambda\eta)\right]_{\lambda=0},$$

if this quantity exists, where η is a function with the same properites as y. This variation δF is a functional of y and η, and is linear in η. If δF has a first variation $\delta^2 F(y,\eta_1,\eta_2)$ with respect to y, then F(y) is said to have a second variation, which will be linear in η_2 as well as η_1. If one takes $\eta_1 = \eta_2 = \eta$ it follows that $\delta^2 F(y,\eta,\eta)$ is given by the expression

$$\delta^2 F(y,\eta,\eta) = \left[\frac{d^2}{d\lambda^2} F(y+\lambda\eta)\right]_{\lambda=0},$$

which is homogeneous of the second order in η. In a similar manner the n-th variation $\delta^n F(y,\eta)$ is a homogeneous quantity of degree n in η, and is given by an analogous formula.

The two articles [52, p. 70; 54, p. 1] which were published after Gateaux's death are incomplete in detail, but are very suggestive. Levy's definition [36, p. 51] makes the differential of a functional appear after the manner of Fréchet, as a quantity which is homogeneous of degree one in the increments of the arguments of the functional, and which differs from the change of the functional by a quantity which approaches zero with the modulus of the increment of the argument. The homogeneity condition replaces the linearity property required by Fréchet. Higher differentials could be defined in an analogous manner.

This conception of a differential of a functional has not been fully developed.

Levy in his text on Functional Analysis [36, p. 50] discusses the generality of the definitions of a differential of a functional which were given by various authors. He shows that the definition of Gateaux has greater generality than that of Fréchet, and that each of them is more general than the definition of Volterra.

Another definition of a differential of a functional was given by Bliss in his course on Functions of Lines in 1916. The functional $F(y)$ is said to have a first differential for a fixed y_0 if there exists a linear functional $L(y, \zeta)$, linear in ζ, such that when $\zeta = y - y_0$

$$F(y) - F(y_0) = L(y, \zeta),$$

for every y in a sufficiently small neighborhood of y_0. The differential of $F(y)$ at $y = y_0$ is defined to be the expression $L(y_0, \zeta)$ for arbitrary ζ. A functional $F(y)$ is similarly said to have a second differential if there exists a functional $B(y, \zeta, \zeta_1)$ bilinear in ζ and ζ_1, such that when $\zeta = \zeta_1 = y - y_0$

$$F(y) = F(y_0) + L(y_0, \zeta) + B(y, \zeta, \zeta)$$

The functional $B(y_0, \zeta, \zeta)$ for arbitrary ζ is called the second differential of $F(y)$. Differentials of higher order could be defined similarly. The expressions so obtained for $F(y)$ are analogous to those obtained for an ordinary function

of several real variables by applying Taylor's formula with integral form of the remainder term.

CHAPTER III

APPLICATIONS OF FUNCTIONALS TO THE CALCULUS OF VARIATIONS

Introduction. The progress which has been made in the application of the theory of functionals to generalizations of the problems of the calculus of variations is only moderate. Volterra [17, p. 48] found, for the problem of minimizing a functional, an analogue of the Euler equation, which for special functionals led to integral or integro-differential equations. Hadamard [59, p. 281] devotes an interesting and suggestive chapter of his Lecons sur le Calcul des Variations to functions of lines and their differentials, but he does not develop any theory of maxima and minima of functionals in general. The same may be said of the short appendix on the calculus of functionals in Vivanti's Elementi del Calcolo delle Variazioni. Fischer and Le Stourgeon made serious efforts to bridge the gap between the two fields. Hahn studied in some detail the problem of Lagrange from the functional standpoint and found an analogue of the Lagrange multiplier rule. Lamson has shown the existence of solutions of functional equations which have an application to this problem, and his theorems have been greatly generalized by Hildebrandt and Graves.

Tonelli studied the integrals of problems of the calculus of variations in the plane from the functional standpoint and shows the importance of the notion of semi-continuity of a functional in the proof of the existence of a minimizing curve.

These are the topics considered in this chapter, except that at the end of the chapter a section on the applications of functions of lines to the Hamilton-Jacobi theory is included, in which the work of Volterra, Fréchet, and Prange is mentioned.

16. A first group of attempts to correlate the theory of functionals and the calculus of variations. A number of writers have tried to connect the theory of functionals and the theory of the calculus of variations, but most of them did no more than to find a first differential and to say that a necessary condition for a functional to have a minimum is that this first differential vanish.

Volterra, in his texts on Functions of Lines [17, p. 50] and Integral Equations [18, p. 23], deduces the first variation of a functional (see page 9), sets it equal to zero and calls the result a first necessary condition for a minimum. For the particular case when

$$F[y] = \int_a^b f[x, y(x)]\,dx + \int_a^b \int_a^b f(x, z)\, y(x)\, y(z)\, dx\, dz,$$

his first variation has the form

$$\delta F = \int_a^b \delta f(x)\, dx \left\{ y(x) + \int_a^b [f(x, z) + f(z, x)]\, y(x)\, dz \right\},$$

and in order to make δF vanish, he sets

$$y(x) + \int_a^b [f(x, z) + f(z, x)]\, y(x)\, dx = 0$$

The problem of finding the extremals for this example therefore leads to the solution of an integral equation. For other special forms of the functional F, the problem of finding the extremals leads to integro-differential equations. After arriving at these integral and integro-differential equations, Volterra, in these books, drops the study of the calculus of variations and devotes himself to the solution of these equations.

Hadamard [59, p. 281] in his treatise on the calculus of variations devotes a chapter to functions of lines. He develops a differential δF after the manner of Volterra, (see page 9), and speaks of its properties, but he does not actually apply his results to problems of maxima and minima of functionals.

Levy, [36, p. 80] in his text on Functional Analysis states that a functional F will have a minimum if $\delta F = 0$, and $\delta^2 F(\eta, \eta) > 0$, where δF, $\delta^2 F$ are differentials in the sense of Fréchet (see page 50). This is an obvious extension to functionals of the theory of maxima and minima from ordinary calculus. He does not pursue the matter further.

An attempt to apply the theory of functionals to the calculus of variations was also made by Vivanti [74, p. 253]. He defined a derivative of a functional $F = \int f(x,y,y')dx$ in the Volterra manner, and expressed its first variation in the customary form (see page 9). With these preliminaries, he states that an absolute minimum will exist if for all

points of a field, and for all values of x the first variation vanishes, if $F_{y'y'} > 0$, and if the integral to be minimized is lower semi-continuous. His analysis of the situation is very brief, and appears in his text as an appendix, indicating a phase of the subject which needs further development.

17. <u>The work of Le Stourgeon</u>. The first person to make a systematic study of the relation of the Fréchet differential to the problems of the calculus of variations was Le Stourgeon [53, p. 357]. Her inspiration was probably obtained from a course on Functions of Lines given by Professor Bliss at the University of Chicago in 1916. However, her article is the first in which such ideas appeared in print.

The definitions of a first and second differential of a functional as given by Le Stourgeon appear in Section 14. The functional F(y) is said to have a minimum at y_0 in the class of functions Y if there exists a neighborhood of order one of y_0 in which $F(y) \geq F(y_0)$ for every arc y of the class Y. With this conception of a minimum, Le Stourgeon was able to deduce certain necessary conditions for the existence of the minimum of a functional.

A necessary condition for a functional to have a minimum is that the first differential of the functional vanish. For a functional F(y), Le Stourgeon (see page 54) obtains a differential $L(\eta)$ in the form

$$L(\eta) = \int_a^b \eta(x)\, d\alpha(x) + \int_a^b \eta'(x)\, d\beta(x), \tag{17.1}$$

where η is the variation of $y(x)$, and the functions $\alpha(x)$ and $\beta(x)$ were defined in Section 14. Thus, a necessary condition that the functional $F(y)$ have a minimum for $y = y_0$ in the class of arcs y joining the end points of y_0 is that the differential $L(\eta)$, the expression (17.1), shall vanish for every function $\eta(x)$ which is continuous and has continuous first derivatives on ab and such that

$$\eta(a) = \eta(b) = 0 .$$

The first differential of a functional, (17.1), is expressed as the sum of two integrals. A necessary condition deduced from the vanishing of this sum of integrals states that if $F(y)$ is to have a maximum or minimum, then the functions $\alpha(x)$ and $\beta(x)$ occurring in the expression (17.1) for the differential $L(\eta)$ must satisfy a relation of the form

$$\beta(x) - \int_a^x \alpha(x)\,dx = kx + l, \tag{17.2}$$

everywhere on the interval ab, where k and l are constants. If the maximum or minimum is to hold with respect to the values of $F(y)$ for all arcs $y(x)$ of the type described, in a neighborhood of $y_0(x)$, then the additional conditions

$$\begin{aligned} \beta'(a) &= \alpha(a+0) - \alpha(a), \\ \beta'(b) &= \alpha(b+0) - \alpha(b), \end{aligned} \tag{17.3}$$

must be satisfied. These latter conditions for the problems

of the calculus of variations lead to transversality conditions at the ends of the interval ab.

In ordinary calculus a further necessary condition for a function to have a minimum is that its second differential be greater than or equal to zero. The analogous condition for functionals is that the bilinear functional $B(\eta, \eta)$ (see page 54) must satisfy the relation

$$B(\eta, \eta) \geq 0, \tag{17.4}$$

for every η of the type previously described in the statement of the first necessary condition.

For the problems of the calculus of variations the functional F(y) has the form

$$F(y) = \int_a^b f(x, y, y')\, dx. \tag{17.5}$$

If the functions $\alpha(x)$, $\beta(x)$ are defined by the relations

$$\alpha(x) = \int_a^x f_y\, dx, \qquad \beta(x) = \int_a^x f_{y'}\, dx,$$

the first variation as given by equation (17.1) becomes

$$\int_a^x (f_y \eta + f_{y'} \eta')\, dx = 0.$$

This form is a familiar one in the theory of the calculus of variations. The necessary condition given by equation (17.2), for this case, when differentiated with respect to x reduces to

$$f_{y'} - \int_a^x f_y\, dx = K,$$

which when again differentiated leads to the well-known Euler equation

$$f_y - \frac{d}{dx} f_{y'} = 0.$$

The condition (17.3) for this case implies that

$$f_{y'}(a) = f_{y'}(b) = 0.$$

This is a transversality condition. Thus one sees that the necessary conditions, deduced from the first variation, for a functional to have a minimum reduce to well-known restrictions on the minimizing arc in the case of the classical theory of the calculus of variations.

If one attempts to get second order conditions by minimizing the second differential, as has been done by Bliss in the theory of the calculus of variations [75, p. 163], a further study of the second variation must be made. For this purpose the second differential may be expressed in the form

$$\begin{aligned} B(\eta,\eta) = \int_a^b \int_a^b \{\eta(x)\eta(y)\, d_{xy}\, p(x,y) \\ + 2\eta(x)\eta'(y)\, d_{xy}\, q(x,y) + \eta'(x)\eta'(y)\, d_{xy}\, r(x,y)\}, \end{aligned} \tag{17.6}$$

where, corresponding to the notation on page 55

$$q = \frac{1}{2}\left[q'(x,y) + q''(x,y)\right].$$

For this functional $B(\eta,\eta)$, the first variation may be represented by

$$(17.7) \qquad 2\int_a^b \left\{ \zeta(x)\, dw(x) + \zeta'(x)\, dw_1(x) \right\},$$

where ζ, ζ' are the increments of η and η' respectively, and

$$w(x) = \int_a^b \left[\eta(y)\, d_y\, p(x,y) + \eta'(y)\, d_y\, q(x,y) \right],$$

$$w_1(x) = \int_a^b \left[\eta(y)\, d_y\, q(x,y) + \eta'(y)\, d_y\, r(x,y) \right].$$

If one applies to the expression (17.7) for B(η, η) the condition given by (17.2), it is evident that the equation

$$(17.8) \qquad L(\eta, x) = w_1 - \int_a^x w\, dx = kx + l,$$

must be satisfied by all variations η which minimize the second variation in a class of functions η (x) having η (a) = η (b) =0.

A condition on the character of the solutions of equation (17.8) needs to be imposed in order to make sure that the functional B(η, η) is greater than or equal to zero. This condition states that no solution η of equation (17.8) can exist which vanishes at x = a and at a point ξ between a and b, but which is not identically zero on the interval a ξ , and which furthermore is such that $\eta'(\xi) \neq 0$. This is an analogue of the Jacobi condition for classical problems of the calculus of variations.

For the particular functionals of the calculus of variations given by the expression (17.5), when one defines

$$p(x,y) = \int_a^y f_{yy}dx \qquad \text{for } a \leq y \leq x,$$
$$= \int_a^x f_{yy}dy \qquad \text{for } x \leq y \leq b,$$

and q and r by similar expressions, it follows that equation (17.8) becomes

$$\int_a^x \Big\{ \eta(y) \left[f_{yy'}(y) - (x-y) f_{yy}(y) \right] dy$$
$$+ \eta'(y) \left[f_{y'y'} - (x-y) f_{yy'}(y) \right] dy \Big\} = Kx + \ell .$$

If this expression is differentiated twice, the resulting equation

$$\eta \left[f'_{yy'} - f_{yy} \right] + \frac{d}{dx} \eta' f_{y'y'} = 0,$$

is the Jacobi equation of the calculus of variations. The necessary conditions expressed in terms of its solutions are consequences of statements in the last paragraph.

Thus, from a purely functional standpoint one is able to deduce necessary conditions for the existence of a minimum of a functional, which when applied to the particular functionals of the calculus of variations lead to some of the classical necessary conditions. Le Stourgeon's work is the first systematic attempt to deduce second order necessary conditions for a minimum of a functional with the help of Fréchet's second differential.

18. The work of Fischer. Fischer [19, p. 383; 20, p. 289; 31, p. 259] made some generalizations of Volterra's derivative of a function of a curve, which were discussed

in Sections 7 and 8. He also applied his results to the problem of minimizing functionals, especially those of the calculus of variations.

In order that a functional L have a minimum it is necessary that its first derivative vanish identically along the minimizing curve [19, p. 289]. If the generalized Volterra derivative of Fischer is used, and the minimum is sought within a restricted class of curves K, as described on page 23, this condition implies that

$$(18.1) \qquad \left[\frac{d\,L(C_\alpha)}{d\alpha}\right]_{\alpha=0} = \int_a^b L'(C,x)\,\eta_\alpha(x,0)\,dx = 0,$$

where the quantities involved were discussed on page 26. As η is an arbitrary function, this implies that

$$(18.2) \qquad L'(C,x) = 0.$$

In the case of the calculus of variations, L(C) may be a definite integral of the form

$$(18.3) \qquad L(C) = \int_a^b f(x,y,y',z_i,z'_i)dx,$$

where C is an arc $y = y(x)$, $(x_1 \leq x \leq x_2)$ of a class K (see page 23). The class K is the totality of all arcs C such that the functions $z_i(x)$ determined by the differential equations

$$z_i' = g_i(x,y,y',z_i),$$

and the initial conditions $z_i(x_1) = z_{i1}$, have prescribed end values $z_i(x_2) = z_{i2}$. The proof is made on the assumption that

the curve C is a normal curve (see page 23). If a function F is defined by the relation

$$F = f + \sum_i \lambda_i (z_i' - g_i),$$

and the set of points ξ_i described on page 24 chosen once and for all, it may be shown that the Volterra derivative, as described on page 26 has the value

$$L'(C, \xi) = \left[F_y(x, y, y', z_i, z_i') - F_{y'}'(x, y, y', z_i, z_i')\right]_{x=\xi}, \tag{18.4}$$

where according to Fischer the functions λ_i are determined as solutions of the equations

$$F_{z_j} - F_{z_j'}' = -\lambda_j' - \sum_i \frac{\partial g_i}{\partial z_j} \lambda_i + f_{z_j} - f_{z_j'}' = 0, \tag{18.5}$$

whose constants of integration are determined by the relation

$$\begin{aligned} &F_y - F_{y'}'\Big|_{x=\xi_j} = \\ &\left[f_y' - f_{y'}' - \sum_i \left\{\lambda_i \frac{\partial g_i}{\partial y} - \frac{d}{dx}\left(\lambda_i \frac{\partial g_i}{\partial y'}\right)\right\}\right]_{x=\xi_j} = 0. \end{aligned} \tag{18.6}$$

Thus, a first necessary condition that the integral (18.3) have a minimum is that the expression (18.4) vanish identically. This expression has the same form as the Euler equation for the Lagrange problem of the calculus of variations. The derivative (18.4) is not uniquely determined since it depends on the arguments ξ_i as well as on C and x. However, the constants of integration occurring in the functions $\lambda_i(x)$ are uniquely determined by equations (18.6) [19, p. 394].

Therefore it follows that there is at most one set of constants which make the expression (18.4) identically vanish, which implies that the derivative (18.4) is independent of the arguments ξ_i. Other methods of determining the constants of integration would lead to similar results.

A definition of the derivative of a function of a surface was formulated by Fischer [20, p. 291] and was discussed in Section 8. A necessary condition for a function of a surface to have a minimum is that its first derivative vanish. If the derivative is expressed in the form (8.1) on page 30 it follows that this condition implies that

$$L'(S,x,y) = 0.$$

In particular, if L is the double integral

$$L(S) = \int\int_R f(x,y,z,p,q)dy\,dx,$$

where p and q are the partial derivatives of z with respect to x and y, Fischer shows that the equation $L' = 0$ is equivalent to the Lagrange equation

$$f_z - \frac{\partial f_p}{\partial x} - \frac{\partial f_q}{\partial y} = 0.$$

Fischer also defined the derivative of a function of a surface relative to a set of surfaces K [20, p. 300]. As was shown on page 31 the additional functions M which characterize these surfaces, do not essentially change the form of the definition of the derivative, but do assist in its evalu-

ation in particular cases. This type of derivative occurs in the theory of the calculus of variations in the study of the Lagrange problem.

The particular problem to which Fischer applies his theory is that of minimizing the double integral L(S) above, in a class of surfaces S subject to the conditions that

$$M_j(S) = \iint_R g_j(x, y, z, p, q)\, dy\, dx = m_j \quad (j = 1, \cdots, m).$$

For this particular problem the derivative of L(S) with respect to the restricted set of surfaces (see equation 8.4) is given by the equation

$$L'(S; x, y, x_1, y_1) =$$

$$f_y - \frac{\partial f_p}{\partial x} - \frac{\partial f_q}{\partial y} + \sum_{j=1}^{m} \lambda_j \left[g_{jz} - \frac{\partial g_{jp}}{\partial x} - \frac{\partial g_{jq}}{\partial y} \right],$$

and a first necessary condition for a minimum is that the right hand side of the last equation shall be equal to zero. The functions λ_j; for this case may be determined by the method indicated on page 32.

For functions of surfaces which have exceptional points or curves, Fischer [31, p. 260] deduced derivatives in the forms (8.7) and (8.9) of Section 8. In order that such functionals have a minimum, it is necessary that L'(S,x,y) vanish and also that

$$a_{ij} = 0,$$

for the case of exceptional points, where the a_{ij} are the

quantities defined on page 34. For the case of exceptional curves, he obtained the auxiliary conditions

$$\sum_{0}^{m} \int_{a}^{l} a_j(s) \frac{\partial^j w}{\partial \eta^j} ds = 0$$

with the help of formula (8.9) on page 36.

Fischer [31, p. 262] obtains an analogue of the Jacobi condition of the calculus of variations for the problem of minimizing a function of a surface with an exceptional point of the second kind (see page 34) from a study of its second derivative. The second derivative is given in the form of equation (8.8) on page 35. If L(S) is a double integral and its second derivative $L''(S,x_0,y_0,x,y)$ vanishes, the second variation, as given on page 35 becomes

$$\left[\frac{d^2L}{d\alpha^2}\right]_{\alpha=0} = \iint \sum_{i,j=1}^{m} a_{ij}(x,y) \frac{\partial^{i+j} w_\alpha(x,y,0)}{\partial x^i \partial y^j} w_\alpha(x,y,0)\, dx\, dy.$$

The analogue of the Jacobi condition is obtained by setting the integrand of the above expression equal to zero. A detailed study of the solutions of the resulting equation has not been made.

19. <u>The theory of functionals and the problem of Lagrange</u>. A few authors have applied the theory of functions of lines to generalizations of the Lagrange problem. Their main result has been to obtain for minimum problems involving functionals an analogue of the Lagrange multiplier rule.

Hahn [55, p. 538] has given the most complete discussion of such problems. He considers a set of r operators

$$\overline{W}_\rho (y_i, a, b) \qquad (\rho = 1, \cdots, r;\ i = 1, \cdots n),$$

which have a unique real value on each arc

(19.1) $\qquad y_i = y_i(x) \qquad (a \leq x \leq b,\ i = 1, \cdots, n),$

of a suitably selected class. Further, he considers m other operators

$$\Phi_\mu (y_i, a, b, t) \qquad (\mu = 1, \cdots, m),$$

which depend also on t on the interval $t_1 \leq t \leq t_2$. His problem is to find in a class of arcs (19.1) which satisfy the relations

$$\overline{W}_\rho (y_i, a, b) = 0, \qquad (\rho = 2, \cdots, r)$$

$$\Phi_\mu (y_i, a, b, t) = 0, \qquad (t_1 \leq t \leq t_2;\ \mu = 1, \cdots, m),$$

one which will minimize the functional W_1.

In order to deduce properties of a minimizing arc $y_i^0(x), a^0, b^0$, he considers variables y_i and parameters a and b defined by the relations

(19.2)
$$y_i = y_i^0(x) + \sum_{l=1}^{K} \varepsilon_l \eta_{il}(x) + [\varepsilon],$$
$$a = a^0 + \sum_{l=1}^{K} \varepsilon_l \alpha_l + [\varepsilon];\ b = b^0 + \sum_{l=1}^{K} \varepsilon_l \beta_l + [\varepsilon],$$

where $[\varepsilon]$ is a quantity tending towards zero with $[\varepsilon_1^2 + \cdots + \varepsilon_K^2]^{1/2}$. For such values of y_i, a, b, the functionals W_i and Φ_μ are supposed to be expressible in forms

$$\overline{W}_\rho = \overline{W}_\rho(y^\circ, a^\circ, b^\circ) + \sum_{\ell=1}^{\kappa} \varepsilon_\ell \overline{V}_\rho(\eta_\ell, \alpha_\ell, \beta_\ell) + [\varepsilon],$$

$$\Phi_\mu(y, a, b, t) = \Phi_\mu(y^\circ, a^\circ, b^\circ, t) + \sum_{\ell=1}^{\kappa} \varepsilon_\ell \Psi_\mu(\eta_\ell, \alpha_\ell, \beta_\ell, t) + [\varepsilon],$$

where both the functions $V_\rho(\eta, \alpha, \beta)$ and $\Psi_\mu(\eta, \alpha, \beta, t)$ are linear and continuous with respect to $\eta_i, \eta_i', \alpha, \beta$.

Furthermore he assumes that for every system of k sets of the quantities

$$\eta_{i\ell}(x), \ \alpha_\ell, \ \beta_\ell, \qquad (\ell = 1, \cdots, \kappa),$$

with suitable continuity properties which satisfy the equations

$$\Psi_\mu(\eta, \alpha, \beta, t) = 0 \qquad (t_1 \leq t \leq t_2;\ \mu = 1, \cdots, m), \tag{19.3}$$

there exists a solution **y(x),a,b** of the form (19.2) which satisfies the equations

$$\Phi_\mu(y, a, b, t) = \Phi_\mu(y^\circ, a^\circ, b^\circ, t) \qquad (t_1 \leq t \leq t_2;\ \mu = 1, \cdots, m).$$

Under these hypotheses a necessary condition for W_1 to have a minimum is that the determinant

$$\left| \overline{V}_\rho(\eta_\sigma, \alpha_\sigma, \beta_\sigma) \right| \qquad (\rho, \sigma = 1, \cdots, r),$$

vanish whenever the arguments $\eta_\sigma, \alpha_\sigma, \beta_\sigma$ of V_ρ satisfy the system of equations (19.3).

An immediate consequence of this statement is that there exist r numbers $g_1, \cdots, g_r$ which are not all zero, such that the expression

$$\overline{V}(\eta, \alpha, \beta) = \sum_{\rho=1}^{r} g_\rho \overline{V}_\rho(\eta, \alpha, \beta), \tag{19.4}$$

vanishes for all continuous differentiable solutions of equation (19.3)

If η_v^*, η_v^{**} $(v = 1, \cdots, n)$ are two sets of functions which with $\alpha^*, \beta^*, \alpha^{**}, \beta^{**}$ satisfy the relations

$$\Psi_\mu(\eta^*, \alpha^*, \beta^*, t) = \Psi_\mu(\eta^{**}, \alpha^{**}, \beta^{**}, t),$$

then on account of the linearity of the functionals Ψ_μ

$$\Psi_\mu(\eta^* - \eta^{**}, \alpha^* - \alpha^{**}, \beta^* - \beta^{**}, t) = 0 .$$

Hence, $\eta^* - \eta^{**}, \alpha^* - \alpha^{**}, \beta^* - \beta^{**}$ satisfy equation (19.3), and therefore the equation

$$V(\eta^* - \eta^{**}, \alpha^* - \alpha^{**}, \beta^* - \beta^{**}) = 0$$

Using the linearity property of V, one finds that

$$V(\eta^*, \alpha^*, \beta^*) = V(\eta^{**}, \alpha^{**}, \beta^{**})$$

Thus one sees that any two sets of values of η, α, β which give to the functionals Ψ_μ equal values, likewise give the function $V(\eta, \alpha, \beta)$ equal values. If one defines

$$\psi_\mu(t) = \Psi_\mu(\eta, \alpha, \beta, t),$$

it follows that $V(\eta, \alpha, \beta)$ depends only on m functions $\psi_\mu(t)$ and therefore

(19.5) $$V(\eta, \alpha, \beta) = U(\psi_1, \cdots, \psi_m) \qquad (m = 1, \cdots \mu)$$

The further assumptions are made that to each $\delta > 0$ there is a $\zeta > 0$ such that for each set of functions $\psi_\mu(t)$

which are continuous on the interval $t_1 \leq t \leq t_2$ and less in absolute value than ζ, there is at least one set of functions $\eta_v(x)$ with continuous derivatives on a^0, b^0 which with two numbers α, β satisfy the relations

$$
\begin{aligned}
&\Psi_\mu(\eta, \alpha, \beta, t) = \psi_\mu(t) \qquad (\mu = 1, \cdot \cdot, m), \\
(19.6) \qquad &\left| \eta_v(x) \right| < \delta, \left| \eta_v'(x) \right| < \delta, \left| \alpha \right| < \delta, \left| \beta \right| < \delta .
\end{aligned}
$$

It may now be shown that the function U defined by equation (19.5) is a linear continuous functional, which may therefore be expressed by a generalization of the Riesz theorem (see page 48) in the form

$$\overline{U}(\psi_1, \cdot \cdot, \psi_m) = - \sum_{\mu=1}^{m} \int_{t_1}^{t_2} \psi_\mu(t)\, d\, G_\mu(t),$$

where the m functions G_μ are of limited variation, and the integral is taken in the sense of Stieltjes. The negative sign in front of the integral is chosen for convenience. On substituting for U its value from equations (19.5) and (19.4), and for $\psi_\mu(t)$ the expression (19.6), one is led to the following analogue of the Lagrange multiplier rule. If the arc $y_i^0(x)$ on the interval a^0, b^0 is to furnish a minimum for the problem stated at the beginning of the section, it is necessary that there exist r numbers $g_1, \cdots, g_r$, not all zero, and m functions of limited variation $H_\mu(t)$ such that the relation

$$\sum_{\rho=1}^{r} g_\rho \nabla_\rho(\eta, \alpha, \beta) + \int_{t_1}^{t_2} \sum_{1}^{m} \Psi_\mu(\eta, \alpha, \beta, t)\, d\, H_\mu(t) \quad o,$$

holds for all numbers α, β and functions η which are continuous and have continuous first derivatives on a^0b^0. This is the analogue of the Lagrange multiplier rule for a minimum problem involving functionals of the type described above. Hahn treats the Mayer problem with variable endpoints as formulated by Bliss [69, p. 305] as a special case, deducing the multiplier rule and end conditions which are given in Bliss' memoir. He also displays the special form his functionals W_ρ must have in order that his general problem may reduce to the problem formulated by Bolza [65, p. 430]. The problem of Lagrange and the so-called isoperimetric problems of the calculus of variations are special cases of these.

Fischer [19, p. 392; 20, p. 289] in his studies of the generalized Volterra derivative has deduced certain results for the Lagrange problem. His results are obtained from a consideration of the derivatives of functionals with respect to a restricted set of curves, and have been mentioned in the last section.

Lamson [70, p. 243] studied the solutions of a particular type of functional equation, and showed that the differential equations of the Lagrange problem are equivalent to a single functional equation of the type he considered. Thus, from a study of the solutions of a general functional equation he is able to show the existence of a solution of the system of differential equations occurring in the problem of Lagrange.

The particular functional equation which Lamson

considered is

$$G\left[y;p\right] = z(p) , \tag{19.7}$$

where p is a variable with a range P of a very general type. He showed that this equation for y has a unique solution $y = Y(z,p)$ which is real single-valued on a suitably restricted range of z, and is continuous as a functional of z. For the Lagrange problem one is expected to minimize the integral

$$\int_a^b f(x,y_i,y_i')dx \qquad (i = 1,\cdots,n),$$

subject to the side conditions

$$\varphi_\alpha(x,y_i,y_i') = 0 \qquad (\alpha = 1,\cdots,m),$$

and the end conditions

$$y_i(a) - h_i = 0, \qquad y_i(b) - k_i = 0.$$

To the equations $\varphi_\alpha = 0$ there may be adjoined the equations

$$\varphi_r(x,y_i,y_i') = z_r(x) \qquad (r = m+1,\cdots,n), \tag{19.8}$$

where $z_i = 0$, for $r \leqq m$. Lamson shows that the equation (19.8) and the end conditions at $x = a$ are equivalent to the single equation

$$\sum_{j=0}^{m} \frac{\partial \varphi_i}{\partial y_j}\bigg|_a \left[y_j(a) - h_j\right] + \int_a^x \varphi_i\,dx = Z_i(x).$$

This equation has the form (19.7) where p denotes a pair (i,x), and the range P is the totality of such sets for $i = 1,2,\cdots,n$,

$a \leq x \leq b$. From the general theory of such equations, one sees that the above equation has a unique solution with properties of the type previously described.

Graves and Hildebrandt [79, p. 127; 80, p. 514] studied implicit functions and differential equations in General Analysis. Their results apply to equations of the form $G(x,y) = 0$ where x and y are points of abstract spaces of the type discussed by Fréchet, and their conclusions are therefore applicable to a very great variety of special cases including those of Lamson.

20. The relation of the Tonelli method to functions of lines. In 1923 Tonelli published a treatise [72] on the calculus of variations in the preface of which he states that he will consider the theory from the standpoint of functions of lines. His purpose is to develop the theory of the calculus of variations with the aid of the idea of semi-continuity. Tonelli's ideas were first presented in a series of articles [60, p. 229; 61, p. 297; 62, p. 448; 63, p. 554; 64, p. 132; 66, p. 49; 67, p. 28; 68, p. 233; 71, p. 167] which furnish a portion of the material of his Fondamenti di Calcolo delle Variazioni.

The classical theory of the calculus of variations was developed with a great deal of aid from the theory of differential equations. Tonelli's aim is to replace existence theorems for differential equations by existence theorems for minimizing curves of integrals. Thus, a theory of the relations of an integral to a class of curves is developed, a

theory of functions of lines. Throughout his work the concept of the semi-continuity of a functional is regarded as a fundamental one.

The first volume of Tonelli's text is divided into three parts, in the first of which is presented a theory of sets of curves. In this section occur theorems which insure the existence in a class of curves of a limiting curve. The second part deals primarily with integrals in parametric form, such as occur in problems of the calculus of variations in the plane. Necessary and sufficient conditions are here developed which will insure that such an integral will have lower semi-continuity. In the third portion of this volume non-parametric integrals are considered and results are obtained similar to those deduced for parametric integrals. This is the preliminary material which is necessary for a study of the calculus of variations from this functional viewpoint.

In order to describe satisfactorily the results of Tonelli it is necessary to give some of his definitions [77, p. 514].

1) If for the parametric integral

$$I(C) = I_C = \int_C F\,(x(t), y(t), x'(t), y'(t))dt, \tag{20.1}$$

the function F_1 defined by the relations

$$F_{x'x'} = -y'^2F_1; \qquad F_{x'y'} = -x'y'F_1; \qquad F_{y'y'} = -x'^2F_1,$$

is greater than or equal to zero for (x,y) in a region A, and $x'^2 + y'^2 \neq 0$, then the integral I_C is called positively

quasi-regular.

2) An ordinary curve is one which is rectifiable, and lies entirely in a region A where the integrand of I has suitable continuity properties.

3) A class of ordinary curves K is said to be complete if every rectifiable accumulation curve of the set is in the set.

Using these definitions, Tonelli derives conditions which insure that an integral of the form (20.1) shall be semi-continuous, among which is the following:

If an integral I_C (20.1) is positively quasi-regular for all ordinary curves of bounded length of a class K, then I_C is lower semi-continuous [72, V. 1, p. 292].

The second volume of Tonelli's treatise deals with the theory of the calculus of variations. He discusses the parametric, ordinary and isoperimetric problems in the plane. For each of these he shows first under what conditions a curve will exist which will give to the integral under consideration an absolute maximum or minimum value. He then discusses the properties of such curves, arriving at among other consequences, the four classical necessary conditions of the calculus of variations. The main emphasis, however, seems to be placed on the existence theorems which are deduced for a great variety of problems with the aid of the notion of semi-continuity. One of the principal existence theorems is given in the following paragraphs.

Theorem. If the integral I_C is positively quasi-

regular, and if for a complete class K of ordinary curves all contained in a bounded portion of a region A, it is possible to determine a function $\Phi(\alpha)$, defined and continuous for all real values of α, and always non-negative and non-decreasing, such that for all curves C of the class K

$$L \le \Phi(I_c),$$

where L is the length of the curve C, then there exists a curve in K which will give to the integral I_C its absolute minimum value [72, V. 2, p. 5].

It will be of interest in indicating Tonelli's method, to outlines the proof of the last theorem. Let $\{C_n\}$ n = 1,2,··· be a sequence of sets of ordinary curves of K, such that for each C_n

$$I_{C_n} < i + \frac{1}{n},$$

i being the greatest lower bound of the values of I_C in K, which we will assume to be finite. Then it follows that

$$L_n \le \Phi(I_{C_n}) \le \Phi(i + \tfrac{1}{n}) \le \Phi(i+1), \tag{20.2}$$

which implies that the curves under consideration are of bounded length. From the theory of such sets of curves it is known that the set $\{C_n\}$ has a rectifiable accumulation curve C_0, which can be proved to lie in the region A, and whose length L_0 satisfies the restriction (20.2), and which belongs to the class K.

The function I_C is lower semi-continuous, as it

satisfies the condition given on page 82. Therefore, for every $\varepsilon > 0$ there exists a ρ such that for any curve C_n which lies in a ρ neighborhood of C_0 the inequality

$$I_{C_0} < I_{C_n} + \varepsilon < i + \frac{1}{n} + \varepsilon,$$

holds. Hence one concludes that

$$I_{C_0} \leq i + \varepsilon \quad ,$$

and therefore $I_{C_0} \leq i$. However, C_0 was shown to belong to the set K, and hence $I_{C_0} \geq i$. Therefore, $I_{C_0} = i$, which implies that C_0 is a rectifiable curve giving to the integral. its absolute minimum value.

Tonelli likewise deduces theorems which insure the existence of a minimizing curve for non-parametric integrals and the integrals of the so-called isoperimetric problem. The hypotheses he uses can generally be divided into two parts, the first of which insures the necessary semi-continuity of the integral I_C; the second specifying the region under consideration, and the properties of the curves in that region. One important consequence of these existence theorems is that they imply the existence of solutions of some of the differential equations utilized in the older theory of the calculus of variations.

In the classical treatment of the calculus of variations, a group of necessary conditions for a minimum is derived. These are also obtained by Tonelli, the difference being that they are occassionally rephrased, in order that the concept

of semi-continuity and the modern theories of integration may be utilized.

The last part of the second volume of Tonelli's text is devoted to a study of the problems of relative maxima and minima within a restricted class of curves. The usual criteria which insure the minimizing property of a curve are deduced for the parametric, ordinary and isoperimetric problems, essentially the same as those deduced in the classical theory. Many of the results obtained for such problems of relative maxima and minima are easily derivable from those obtained in the first part of this volume by suitably restricting the class of curves among which the extremum is sought.

Since the publication of the Tonelli treatise a few authors have generalized or extended the ideas which Tonelli introduced, among them being Roussel [78, p. 395], Hahn [76, p. 437], Downing [82], and Mc Shane [83]. Their work has been largely for problems in the plane. Whether these methods of Tonelli are applicable to problems of a more general type such as the Lagrange problem remains to be seen.

21. The Hamilton-Jacobi theory for multiple integrals. The Hamilton-Jacobi theory for multiple integrals, as so far developed involves in an essential manner the theory of functionals. For double or multiple integrals the analogue of the Hamilton-Jacobi partial differential equation is a functional equation. The relations between such equations and the problems of the calculus of variations are at the present time only imperfectly developed. In this section a sketch is

given of the results of Volterra, Fréchet, and Prange.

Volterra [14, p. 127] was the first person to study the Hamilton-Jacobi theory for multiple integrals from a functional standpoint. He considered the problem of minimizing the double integral

(21.1) $$I = \iint F(u, v, x_1, \cdots, x_n, \xi_{ik})\, du\, dv,$$

where

$$\xi_{ik} = \frac{d(x_i, x_k)}{d(u, v)}.$$

The first variation of the integral (21.1) is found to be

$$\delta I = \iint \left(\sum \frac{\partial F}{\partial \xi_{ik}} \delta \xi_{ik} + \sum_1^n \frac{\partial F}{\partial x_i} \delta x_i \right) du\, dv,$$

which must be equal to zero when the integral (21.1) takes on a minimum value. On integrating by parts, and assuming that δx_i vanishes on the boundary of the region over which the integration is performed, one finds that on a minimizing surface

(21.2) $$\frac{\partial F}{\partial x_i} - \sum_{k=1}^n \frac{d\left(\frac{\partial F}{\partial \xi_{ik}}, x_k\right)}{d(u, v)} = 0 \qquad (i = 1, 2, \cdots, n).$$

For convenience put

(21.3) $$\frac{\partial F}{\partial \xi_{ik}} = p_{ik}; \qquad p_{ii} = 0,$$

and

(21.4) $$H = -F + \sum p_{ik} \xi_{ik}.$$

From the relations (21.2) and (21.3) it follows that

$$\sum_{k=1}^{n} \frac{d(p_{ik}, x_k)}{d(u,v)} = \frac{\partial F}{\partial x_i}.$$

On solving equations (21.3) for ξ_{ik} and substituting these values in equation (21.4), H becomes a function of the form

$$H = H(u, v, x_1, \cdots, x_n, p_{ik}).$$

From equation (21.4) the relations

$$(21.5) \qquad \frac{\partial H}{\partial x_i} = -\frac{\partial F}{\partial x_i}, \qquad \frac{\partial H}{\partial p_{ik}} = \xi_{ik},$$

follow, which with the help of the above notation may be expressed as

$$(21.6) \qquad \frac{d(x_i, x_k)}{d(u,v)} = \frac{\partial H}{\partial p_{ik}}, \qquad \sum_{k=1}^{n} \frac{d(p_{ik}, x_k)}{d(u,v)} = -\frac{\partial H}{\partial x_i}.$$

This form of the equations is perfectly analogous to the canonical form given for the equations of dynamics by Hamilton, and is called by Volterra the canonical system for this problem.

Volterra also shows that if one is given a system of differential equations of the form (21.6), an integral of the calculus of variations will exist whose extremals are determined by these equations. Such an integral is

$$J = \iint \left(\sum p_{ik} \frac{d(x_i, x_k)}{d(u,v)} - H \right) du\, dv.$$

Volterra then studies in some detail the case when there are three dependent variables x_1, x_2, x_3. For this situ-

ation he derives a number of theorems, two of the principal ones being stated here without proof. In these theorems it is assumed that H does not depend explicitly on u and v, and that W is a function of a line depending on the curves in $x_1x_2x_3$-space which correspond to the boundary of the region of integration in the uv-plane. The quantities $\partial W/\partial(x_i, x_k)$ are functional derivatives of a type similar to those used by Prange which are described more in detail later in this section.

Theorem I. If the integrals of the differential equations

$$\frac{d(x_i, x_k)}{d(u,v)} = \frac{\partial H}{\partial p_{ik}}, \quad \sum \frac{d(p_{ik}, x_k)}{d(u,v)} = -\frac{\partial H}{\partial x_i} \quad (i = 1,2,3),$$

are known, it is possible to determine a function W of a line in $x_1x_2x_3$-space of the first degree which satisfies the relation

$$H\left(\frac{\partial \overline{W}}{\partial(x_2,x_3)}, \frac{\partial \overline{W}}{\partial(x_3,x_1)}, \frac{\partial \overline{W}}{\partial(x_1,x_2)}, x_1, x_2, x_3\right) + h = 0,$$

where h is a constant, and the derivatives of the function W are substituted for the quantities p_{ik} in H.

Theorem II. Let W be a function of a line of the first degree, depending on a curve in a three-dimensional space with coordinates x_1, x_2, x_3 and satisfying the relation

$$H\left(\frac{\partial \overline{W}}{\partial(x_2,x_3)}, \frac{\partial \overline{W}}{\partial(x_3,x_1)}, \frac{\partial \overline{W}}{\partial(x_1,x_2)}, x_1, x_2, x_3\right) = h,$$

where h is a constant. Let also p_{23}, p_{31}, p_{12} be defined by the equations

$$\frac{\partial \overline{W}}{\partial(x_2,x_3)} = p_{23}, \qquad \frac{\partial \overline{W}}{\partial(x_3,x_1)} = p_{31}, \qquad \frac{\partial \overline{W}}{\partial(x_1,x_2)} = p_{12}.$$

If on substituting these values in the equations

$$\frac{d(x_i\, x_k)}{d(u,v)} = \frac{\partial H}{\partial p_{ik}},$$

it turns out that they are compatible, then these values also satisfy the equations

$$\sum_k \frac{\partial(p_{ik}, x_k)}{\partial(u,v)} = -\frac{\partial H}{\partial x_i}.$$

This last theorem implies the existence of an extremal surface for the original problem because of the relations existing between the so-called canonical equations and the equations for the extremals of the integral.

Fréchet [58, p. 187] extended Volterra's work to the case when one has r independent variables. He deduced results analogous to Volterra's for an integral of the form

$$I = \iint\cdots\int f\left(x_{r+1},\cdots,x_n; x_1,\cdots,x_r,\ \frac{\partial x_{r+1}}{\partial x_1},\cdots,\frac{\partial x_n}{\partial x_r}\right) dx_1\cdots dx_r,$$

where $x_1,\cdots,x_r$ are the independent variables, $x_{r+1},\cdots,x_n$, the dependent ones, and the partial derivatives $\partial x_{r+k}/\partial x_i$ are expressible in the form

$$\frac{\partial x_{r+k}}{\partial x_i} = \frac{D(x_1,\cdots x_{i-1}, x_{r+k}, x_{i+1},\cdots,x_r)}{D(w_1,\cdots,w_r)} : \frac{D(x_1,\cdots,x_r)}{D(w_1,\cdots,w_r)}$$

Prange [30] likewise discussed the Hamilton-Jacobi theory for multiple integrals. He showed, as in Section 4, that the differential of a functional S [L] , where L is a space curve, is expressible in the form

$$\delta S[L] = \int_L \left(S_x \delta x + S_y \delta y + S_z \delta z\right) d\sigma,$$

where $d\sigma$ is the arc length along L. Prange states [30, p. 40] that if the curve L is displaced along itself then

$$S_x \frac{dx}{d\sigma} + S_y \frac{dy}{d\sigma} + S_z \frac{dz}{d\sigma} = 0,$$

which may be shown to be equivalent to the fact that S [L] is by hypothesis independent of the parametric representation of L. The vector (S_x, S_y, S_z) is therefore perpendicular to the curve L, and its components may be expressed as the determinants of a matrix of the form

$$\left\| \begin{matrix} \frac{dx}{d\sigma}, & \frac{dy}{d\sigma}, & \frac{dz}{d\sigma} \\ S_{yz}, & S_{zx}, & S_{xy} \end{matrix} \right\| .$$

The quantities S_{yz}, S_{zx}, S_{xy} are called by Volterra [6, p. 278] and Prange [30, p. 41] the functional derivatives of S with respect to the coordinate planes. These quantities are not uniquely determined, for they would satisfy the same relations if an arbitrary multiple of $dx/d\sigma$,$dy/d\sigma$,$dz/d\sigma$ were added to each of them respectively. In terms of them it follows that δS may be represented in the form

$$\delta S = -\int_L \begin{vmatrix} \delta x, & \delta y, & \delta z \\ \frac{dx}{d\sigma}, & \frac{dy}{d\sigma}, & \frac{dz}{d\sigma} \\ S_{yz}, & S_{zx}, & S_{xy} \end{vmatrix} d\sigma, \tag{21.7}$$

The double integral

$$I = \iint_S f(x,y,z,p,q)\,dx\,dy,$$

where $p = dz/dx$, $q = dz/dy$, is a function depending on the surface S bounded by a space curve L. If a fixed, simply closed curve in space L_0, is given, Prange states that through L_0 and a neighboring curve L a unique extremal surface E of the integral may be passed. Thus, the integral I taken on such a surface depends on the two curves L_0 and L, and can be considered as a function of the variable curve L. For such a functional it may be shown [30, p. 44] that

$$\delta I = -\int_L \begin{vmatrix} \delta x, & \delta y, & \delta z \\ \frac{dx}{d\sigma}, & \frac{dy}{d\sigma}, & \frac{dz}{d\sigma} \\ f_p, & f_q, & pf_p + qf_q - f \end{vmatrix} d\sigma, \tag{21.8}$$

where $p:q:-1$ is the direction of the normal to the extremal surface E at a point of L. After comparing this expression with the expression (21.7), Prange states that by comparison of equations (21.7) and (21.8) he is justified in setting

$$I_{yz} = f_p, \quad I_{zx} = f_q, \quad I_{xy} = pf_p + qf_q - f. \tag{21.9}$$

If one introduces the canonical variables

(21.10) $\pi = f_p, \qquad k = f_q,$

which imply on solving for p and q that

(21.11) $\frac{dz}{dx} = p = P(x,y,z,\pi,k), \qquad \frac{dy}{dx} = q = Q(x,y,z,\pi,k),$

and if one defines

(21.12) $H(x,y,z,\pi,k) = pf_p + qf_q - f \Big|_{p = P,\ q = Q},$

then according to Prange the analogue of the Hamilton-Jacobi partial differential equation is the functional equation

(21.13) $I_{xy} = H(x,y,z,I_{yz},I_{zx}),$

which is a consequence of equation (21.9). The problem is to find a functional I, satisfying equation (21.13), and whose functional derivatives I_{xy}, I_{yz}, I_{zx} satisfy the relations (21.9) and (21.12). If such a functional is found, equations (21.11) would determine a set of extremal surfaces.

Caratheodory [73, p. 78; 81, p. 193] also considered the problem of finding canonical equations for the extremals of problems in multiple integrals of the calculus of variations. He did not, however, explicitly use the idea of a function of a surface.

The problem of finding solutions of a set of functional canonical equations such as (21.11) and (21.12), and the properties of such solutions, such as the transversality configurations and the determination of fields have not been fully developed. The studies which have been made in this domain are very suggestive. There remain in this field many interesting questions which have not been completely answered.

BIBLIOGRAPHY

In this bibliography the following abbreviations are used:

Am. Jnl.	American Journal of Mathematics.
Ann. Sci.	Annales Scientifiques de L'École Normale Supérieure.
Bull. A.M.S.	Bulletin of the American Mathematical Society.
Bull. S.M.F.	Bulletin de la Société Mathématique de France.
Lincei.	Atti della R. Accademia dei Lincei, Rendiconti.
Nouv. Ann.	Nouvelles Annales de Mathématique.
Palermo.	Rendiconti del Circolo Matematico di Palermo.
Trans.	Transactions of the American Mathematical Society.

CHAPTER I

THE VOLTERRA DERIVATIVE

1. Koenigs, Sur les intégrales de certaines équations fonctionnelles. Ann. Sci., Ser. 3, I (1884), p. S. 3.

2. Volterra, Sopra le funzioni che dipendono da altre funzioni, Nota I. Lincei, Ser. 4, III 2 (1887), p. 97.

3. Volterra, Sopra le funzioni che dipendono da altre funzioni, Nota II. Lincei, Ser. 4, III 2 (1887), p. 141.

4. Volterra, Sopra le funzioni che dipendono da altre funzioni, Nota III. Lincei, Ser. 4, III 2 (1887), p. 153.

5. Volterra, Sopra le funzioni dipendenti da linee, Nota I. Lincei, Ser. 4, III 2 (1887), p. 225.

6. Volterra, Sopra le funzioni dipendenti da linee, Nota II. Lincei, Ser. 4, III 2 (1887), p. 274.

7. Volterra, Sopra una estensione della teoria di Riemann sulle funzioni di variabili complesse, Nota I. Lincei, Ser. 4, III 2 (1887), p. 281.

8. Volterra, Sopra una estensione della teoria di Riemann sulle funzioni di variabili complesse, Nota II. Lincei, Ser. 4, IV (1888), p. 107.

9. Volterra, Sopra una estensione della teoria di Riemann sulle funzioni di variabili complesse, Nota III. Lincei, Ser. 4, IV (1888), p. 196.

10. Volterra, Delle variabili complesse negle iperspazi, Nota I. Lincei, Ser. 4, V (1889), p. 158.

11. Volterra, Delle variabili complesse negli iperspazi, Nota II. Lincei, Ser. 4, V (1889), p. 291.

12. Volterra, Sulle funzioni coniugate. Lincei, Ser. 4, V (1889), p. 599.

13. Volterra, Sur une généralisation de la théorie des fonctions d'une variable imaginaire. Acta Mathematica, XII (1889), p. 233.

14. Volterra, Sopra une estensione della teoria Jacobi-Hamilton del calcolo delle variazioni. Lincei, Ser. 4, VI 1 (1890), p. 127.

15. Fabri, Sopra alcune proprietá generali delle funzioni che dipendono da altre funzioni e da linee. Atti della R. Accademia delle Scienze di Torino, XXV (1890), p. 432.

16. Hadamard, Sur les dérivées des fonctions de lignes. Bull. S.M.F. XXX (1902), p. 40.

17. Volterra, Leçons sur les fonctions de lignes. (1913).

18. Volterra, Leçons sur les équations intégrales. (1913).

19. Fischer, A generalization of Volterra's derivative of a function of a curve. Am. Jnl., XXXV (1913), p. 369.

20. Fischer, The derivative of a function of a surface. Am. Jnl., XXXVI (1914), p. 289.

21. Pascal, Sui principii della teoria delle funzioni di linee, Nota I. Rendiconti dell'Accademia delle Scienze Fische e Matematiche di Napoli, Ser. 3, XX (1914), p. 40.

22. Pascal, Sui principii della teoria delle funzioni di linee, Nota II. Rendiconti dell'Accademia delle Scienze Fische e Matematiche di Napoli, Ser. 3, XX (1914), p. 68.

23. Bliss, A note on functions of lines. Proceedings of the National Academy of Sciences, I (1915), p. 173.

24. Daniele, Formole di derivazione funzionale, Nota I. Lincei, Ser. 5, XXIV 1 (1915), p. 319.

25. Daniele, Formole di derivazione funzionale, Nota II. Lincei, Ser. 5, XXIV 1 (1915), p. 496.

26. Daniele, Sulle derivate delle funzioni di linee inverse. Atti della Accademia Gioenia di Scienze Naturali in Catenia, VIII (1915), Memoire XIII.

27. Daniele, Osservazione e teoremi sulle funzioni di linee e loro derivate. Giornali di Matematiche di Battaglini, LIII (1915), p. 162.

28. Evans, Note on the derivative and the variation of a function depending on all the values of another function. Bull. A.M.S., XXI (1915), p. 387.

29. Pascal, La linee funzioni di linee. Giornali di Matematiche di Battaglini, LIII (1915), p. 318.

30. Prange, Die Hamilton-Jacobische Theorie für Doppelintegrale. Göttingen Dissertation, (1915).

31. Fischer, Functions of surfaces with exceptional points or curves. Am. Jnl., XXXVIII (1916) p. 259.

32. Fischer, Equations involving the partial derivatives of a function of a surface. Am. Jnl., XXXIX (1917), p. 123.

33. Kakeya, Theory of analytic functions of lines. Science Reports of the Tohoku Imperial University, Ser. 1, VI (1917), p. 341.

34. Kakeya, On functions of lines and sets of curves. Science Reports of the Tohoku Imperial University, Ser. 1, VII (1918), p. 177.

35. Daniell, The derivative of a functional. Bull. A.M.S., XXV (1919), p. 414.

36. Levy, Leçons D'Analyse Fonctionnelle, (1922).

37. Mandelbrojt, Sulla generalizzazione del calcolo delle variazione. Lincei, Ser. 6, I (1925), p. 151.

38. Volterra, Teoria de las funcionales y de las ecuaciones integrales e integro-diferenciales. (1927).

CHAPTER II

THE FRÉCHET DIFFERENTIAL OF A FUNCTIONAL

39. F. Riesz, Sur les ópérations fonctionnelles linéaires. Comptes Rendus, CXLIX (1909), p. 974.

40. Fréchet, Extension au cas des intégrales multiples d'une definition de l'intégrale due a Stieltjes. Nouv. Ann., Ser. 4, X (1910), p. 241.

41. F. Riesz, Sur certaines systèmes singuliers d'equations intégrales. Ann. Sci., XXVIII (1911), p. 33.

42. Fréchet, Sur la notion de différentiele totale. Nouv. Ann., Ser. 4, XII (1912), p. 385.

43. Gateaux, Sur la représentation des fònctionnelles continues. Lincei, Ser. 5, XXII 2 (1913), p. 646.

44. Fréchet, Sur le notion de différentiele d'une fonction de ligne. Trans., XV (1914), p. 135.

45. Gateaux, Sur la représentation des fonctionnelles continues. Lincei, Ser. 5, XXIII 1 (1914), p. 310.

46. Gateaux, Sur les fonctionnelles d'ordre entier d'approximation. Lincei, Ser. 5, XXIII 1 (1914), p. 405.

47. F. Riesz, Démonstration nouvelle d'un théoreme concernant les ópérations fonctionnelles linéaires. Ann. Sci., XXXI (1914), p. 2.

48. Fréchet, Sur les fonctionnelles bilinéaires. Trans., XVI (1915), p. 215.

49. Fischer, On bilinear and n-linear functionals. Proceedings of the National Academy of Sciences, III (1917), p. 640.

50. Fischer, Linear functionals of n-spreads. Annals of Mathematics, XIX (1917-18), p. 37.

51. Hildebrandt, On integrals related to and extensions of the Lebesgue integrals. Bull. A.M.S., XXIV (1918), p.183.

52. Gateaux, Fonctions d'une infinité de variables indépendantes. Bull. S·M.F., XLVII (1919), p. 70.

53. Le Stourgeon, Minima of functions of lines. Trans., XXI (1920), p. 357.

54. Gateaux, Sur diverses questions de calcul fonctionnel. Bull. S·M.F., L (1922), p. 1.

55. Hahn, Über die Lagrangesche Multiplikatorenmethode. Sitzungsberichte der Akademie Wien, CXXXI (1922), p. 531.

56. Fréchet, La notion de différentiele dans l'analyse générale. Ann. Sci., Ser. 3, XLII (1925), p. 293.

57. Wong, Representation of bilinear functionals in terms of Stieltjes integrals. Master's Dissertation, University of Chicago, (1929).

CHAPTER III

APPLICATIONS OF FUNCTIONALS TO THE CALCULUS OF VARIATIONS

58. Fréchet, Sur une extension de la méthode de Jacobi-Hamilton. Annali di Matematica, Ser. 3, XI (1905), p. 187.

59. Hadamard, Leçons sur le calcul des variations. (1910).

60. Tonelli, Sugli integrali curvilinei del calcolo delle variazioni, Nota I. Lincei, Ser. 5, XX 1 (1911), p. 229.

61. Tonelli, Sui massimi e minimi assoluti del calcolo delle variazioni. Palermo, XXXII (1911), p. 297.

62. Tonelli, Sugli integrali curvilinei del calcolo delle variazioni, Nota II. Lincei, Ser. 5, XXI 1 (1912), p. 448.

63. Tonelli, Sugli integrali curvilinei del calcolo delle variazioni. Nota III. Lincei, Ser. 5, XXI 1 (1912), p. 554.

64. Tonelli, Sugli integrali curvilinei del calcolo delle variazioni, Nota IV. Lincei, Ser. 5, XXI 2 (1912), p. 132.

65. Bolza, Über der "Anormalen Fall" beim Lagrangeschen und Mayerschen Problem mit gemischten Bedingungen und variablen Endpunkten. Mathematische Annalen, LXXIV

(1913), p. 430.

66. Tonelli, Sul caso regolare nel calcolo delle variazioni. Palermo, XXXV (1913), p. 49.

67. Tonelli, Sulle funzioni di linee. Lincei, Ser. 5, XXIII 1 (1914), p. 28.

68. Tonelli, Sur une méthode directe du calcul des variations. Palermo, XXXIX (1915), p. 233.

69. Bliss, The problem of Mayer with variable endpoints. Trans. XIX (1918), p. 305.

70. Lamson, A general implicit function theorem with an application to problems of relative minima. Am. Jnl., XLI (1920), p. 243.

71. Tonelli, La semi-continuitá nel calcolo delle variazioni. Palermo, XLIV (1920), p. 167.

72. Tonelli, Fondamenti di calcolo delle variazioni. Vol. 1, (1921); Vol. 2, (1923).

73. Caratheodory, Über die kanonischen Veränderlichen in der Variationsrechnung der mehrfachen Integrale. Mathematische Annalen, LXXXV (1922), p. 78.

74. Vivanti, Elementi del calcolo delle variazioni. (1923).

75. Bliss, Calculus of variations. (1925).

76. Hahn, Über ein Existenz-theorem der Variationsrechnung. Sitzungsberichte der Akademie Wien, Abt. 2A, CXXXIV (1925-26), p. 437.

77. Dresden, Some recent work in the calculus of variations. Bull. A.M.S., XXXII (1926), p. 475.

78. Roussel, Recherches sur le calcul des variations. Journal de Mathématique, Ser. 9, V (1926), p. 395.

79. Hildebrandt-Graves, Implicit functions and their differentials in general analysis. Trans., XXIX (1927), p. 127.

80. Graves, Implicit functions and differential equations in general analysis. Trans., XXIX (1927), p. 514.

81. Caratheodory, Über der Variationsrechnung bei mehrfachen Integralen. Acta Litterarum ac Scientiarum, IV (1929), p. 193.

82. Downing, Absolute minima for space problems of the calculus of variations in parametric form. University of Chicago, Doctor's Dissertation, (1929).

83. Mc Shane, Semi-continuity in the calculus of variations, and absolute minima for isoperimetric problems. University of Chicago Doctor's Dissertation, (1930).

SUFFICIENT CONDITIONS FOR A PROBLEM OF MAYER IN THE CALCULUS OF VARIATIONS

by

GILBERT AMES BLISS

and

MAGNUS RUDOLPH HESTENES

TABLE OF CONTENTS

SUFFICIENT CONDITIONS FOR A PROBLEM OF MAYER
IN THE CALCULUS OF VARIATIONS[1]

1. Introduction. The general problem of Mayer with variable end-points as proposed by Bliss [V, p. 305][2] is that of finding in a class of arcs

$$(1{:}1)\qquad y_i = y_i(x) \qquad (x_1 \leqq x \leqq x_2;\ i = 1, \cdots, n)$$

satisfying a system of differential equations and end-conditions

$$\phi_\alpha(x, y, y') = 0 \qquad (\alpha = 1, \cdots, m < n),$$
$$\psi_\mu[x_1, y(x_1), x_2, y(x_2)] = 0 \ (\mu = 1, \cdots, p \leqq 2n + 1)$$

one which minimizes a function of the form

$$g[x_1, y(x_1), x_2, y(x_2)] \ .$$

Bliss has shown that this problem is equivalent to a problem of Bolza [V, p. 306] in the sense that each can be transformed into one of the other type. For the problem of Bolza the function to be minimized is

$$I = g[x_1, y(x_1), x_2, y(x_2)] + \int_{x_1}^{x_2} f(x, y, y')dx,$$

[1]It is expected that this paper will appear in the Transactions of the American Mathematical Society, vol. 35.

[2]The numbers in the square brackets refer to the bibliography at the end of the text.

and it is clear at once that the problem of Mayer is a problem of Bolza having $f \equiv 0$.

Sufficient conditions for the problem of Bolza have been established by Morse [XI, p. 528] and Bliss [XII, p. 271]. However the hypotheses which they make, in particular that of normality on every sub-interval, imply that the function f is not identically zero, and the sets of sufficient conditions established by them are therefore not applicable to the problem of Mayer without further modification. In view of this fact it is the purpose of the authors of the present paper to establish a set of sufficient conditions for the problem of Mayer with variable end-points. This will be done in two parts, the first of which is the paper here presented, dealing only with the special case in which the number of end-conditions $\psi_\mu = 0$ is exactly $2n + 1$. By methods similar to those used by Bliss for the problem of Bolza [XII, p. 261 - 274] the results obtained will be extended to the general case in a second paper by Hestenes.

The problem considered here is an obvious generalization of the classical problem of Mayer and reduces to the latter when the expression to be minimized is the function $g = y_1(x_2)$ and the end-conditions $\psi_\mu = 0$ are the conditions

$$x_1 - \alpha_1 = y_i(x_1) - \beta_{i1} = x_2 - \alpha_2 = y_j(x_2) - \beta_{j2} = 0 \qquad (i = 1, \cdots, n;\ j = 2, \cdots, n),$$

the α's and β's being constants. Sufficiency theorems for

the classical problem have been established by Egorov [II, p. 376], Kneser [I, p. 250; VIII, p. 290], and Larew [VII, p. 65], who use in each case an n-dimensional field defined in the $(n + 1)$-dimensional space of points $(x, y_1, \cdots, y_n)$ by an $(n - 1)$-parameter family of extremals passing through a fixed point. Such a field does not seem to be applicable to the problem considered here, but one can use instead a field of $n + 1$ dimensions defined by an n-parameter family of extremals in $(x, y_1, \cdots, y_n)$-space. The construction and use of such a field are important features of this paper. An $(n + 1)$-dimensional field of this sort is applicable to the more special classical problem of Mayer also, and a fundamental sufficiency theorem for this case can be established in this way with greater ease and fewer restrictions than have hitherto been required.

2. Preliminary remarks. In the following pages it is assumed that the various indices have the following ranges unless otherwise explicitly specified:

$$i,\ k = 1, 2, \cdots, n \quad ; \alpha, \beta = 1, 2, \cdots, m < n;$$
$$\rho,\ \sigma = 1, 2, \cdots, 2n + 1; \qquad r = 1, 2, \cdots, n - 1;$$
$$s = 1, 2, \cdots, 2n - 1.$$

The tensor analysis summation convention is used freely throughout. We make the following hypotheses concerning a particular arc E_{12} whose minimizing properties are to be studied:

(a) The functions $y_i(x)$ defining E_{12} are continuous on the interval x_1x_2, and this interval can be subdivided into a finite number of parts on each of which these functions have continuous derivatives.

(b) The functions ϕ_α have continuous partial derivatives of the first three orders in a neighborhood $\mathfrak{R}$ of the values (x, y, y') on E_{12}, and at each element (x, y, y') in $\mathfrak{R}$ the matrix $\| \phi_{\alpha y_k'} \|$ has rank m.

(c) The functions g, ψ_ρ have continuous partial derivatives of the first two orders in a neighborhood of the end-values $(x_1, y_{i1}, x_2, y_{i2})$ of E_{12} in which the determinant

$$(2:1) \qquad \begin{vmatrix} g_{x_1} & g_{y_{i1}} & g_{x_2} & g_{y_{i2}} \\ \psi_{\rho x_1} & \psi_{\rho y_{i1}} & \psi_{\rho x_2} & \psi_{\rho y_{i2}} \end{vmatrix}$$

is different from zero.

An <u>admissible set</u> (x, y, y') is a set interior to $\mathfrak{R}$ and satisfying the equations $\phi_\alpha = 0$. An arc (1:1) having the continuity properties described in (a) is called <u>admissible</u> if all of its elements (x, y, y') are admissible. The definitions of equations of variation and of admissible variations used in the following pages are those of Bliss [V, p. 307; IX, p. 677]. The problem of Mayer here proposed can now be more precisely stated as that of finding in the class of admissible arcs satisfying the end-conditions $\psi_\rho = 0$ one which minimizes the function g.

I. THE FIRST NECESSARY CONDITION. For every minimizing arc E_{12} for the problem of Mayer as here proposed there exist constants c_i and a function $F = \lambda_\alpha(x)\,\phi_\alpha$ such that the equations

$$(2:2) \qquad F_{y_i'} = \int_{x_1}^{x} F_{y_i}\,dx + c_i\,, \qquad \phi_\alpha = 0$$

are satisfied at each point of E_{12}. The multipliers $\lambda_\alpha(x)$ are continuous except possibly at the values of x defining corners of E_{12} and do not vanish simultaneously at any point of E_{12}.

To prove this theorem one needs only to combine the methods used by Bliss for the corresponding theorems in the problems of Mayer [V, p. 311] and Lagrange [IX, p. 683]. It is also an immediate corollary of a theorem established by Morse and Myers for the problem of Bolza [X, p. 245].

THEOREM 2:1. If the functions $\lambda_\alpha(x)$ are a set of multipliers with which an admissible arc E_{12} satisfies the equations (2:2), then for every set of admissible variations ξ_1, ξ_2, $\eta_i(x)$ along E_{12} the functions $\eta_i(x)$ satisfy the equations

$$(2:3) \qquad F_{y_i'}\,\eta_i \Big|_{x'}^{x''} = 0$$

for every interval $x'x''$.

This result is readily provable by multiplying the equations of variation

$$\phi_{\alpha y_i}\eta_i + \phi_{\alpha y_i'}\eta_i' = 0$$

by the multipliers $\lambda_\alpha(x)$, adding, and applying the usual integration by parts with the help of equations (2:2).

An admissible arc E_{12} is said to be <u>normal relative to the end-conditions</u> $\psi_\rho = 0$ if there exist for it 2n + 1 sets of admissible variations ξ_1^σ, ξ_2^σ, $\eta_i^\sigma(x)$ such that the determinant $|\Psi_\rho(\xi^\sigma, \eta^\sigma)|$ is different from zero, where

$$\Psi_\rho(\xi, \eta) = (\psi_{\rho x_1} + y_{i1}'\psi_{\rho y_{i1}})\xi_1 + \psi_{\rho y_{i1}}\eta_{i1}$$
$$+ (\psi_{\rho x_2} + y_{i2}'\psi_{\rho y_{i2}})\xi_2 + \psi_{\rho y_{i2}}\eta_{i2},$$

the functions y_i, y_i' occurring explicitly and in the derivatives of ψ_ρ being those belonging to E_{12}. The arc E_{12} is <u>normal on the sub-interval</u> $x'x''$ if there exist for it 2n - 1 sets of admissible variations ξ_1^s, ξ_2^s, $\eta_i^s(x)$ such that the matrix

$$\left\| \begin{matrix} \eta_i^s(x') \\ \eta_i^s(x'') \end{matrix} \right\| \tag{2:4}$$

has rank 2n - 1. On account of the relation (2:3) this is the highest rank attainable for a matrix with columns of this sort belonging to an arc that satisfies the equations (2:2) with a set of multipliers $\lambda_\alpha(x)$. For convenience an arc that is normal relative to the end-conditions $\psi_\rho = 0$ will be designated simply as <u>normal</u>.

THEOREM 2:2. <u>An admissible arc E_{12} that satisfies the necessary condition I is normal if and only if there exist for it no set of multipliers $\lambda_\alpha(x)$, not vanishing simultaneously</u>,

with which it satisfies equations (2:2) and for which the determinant

$$(2:5)\quad \begin{vmatrix} 0 & F_{y_1'}(x_1) & 0 & -F_{y_1'}(x_2) \\ \psi_{\rho x_1} + y_{i1}'\psi_{\rho y_{i1}} & \psi_{\rho y_{i1}} & \psi_{\rho x_2} + y_{i2}'\psi_{\rho y_{i2}} & \psi_{\rho y_{i2}} \end{vmatrix}$$

vanishes on E_{12}. If E_{12} is normal the constant χ_0 defined below can be taken equal to unity, its multipliers $\lambda_\alpha(x)$ being then unique.

To prove the theorem we first notice that the arc E_{12} is normal if and only if there exist for it no set of constants and multipliers χ_0, χ_ρ, $\lambda_\alpha(x)$ having $\chi_0 = 0$ but not vanishing simultaneously, with which it satisfies the relations (2:2) and

$$(2:6)\quad \begin{aligned} &\chi_0(g_{x_1} + y_{i1}'g_{y_{i1}}) + \chi_\rho(\psi_{\rho x_1} + y_{i1}'\psi_{\rho y_{i1}}) = 0, \\ &\chi_0 g_{y_{i1}} + \chi_\rho \psi_{\rho y_{i1}} = F_{y_i'}(x_1), \\ &\chi_0(g_{x_2} + y_{i2}'g_{y_{i2}}) + \chi_\rho(\psi_{\rho x_2} + y_{i2}'\psi_{\rho y_{i2}}) = 0, \\ &\chi_0 g_{y_{i2}} + \chi_\rho \psi_{\rho y_{i2}} = -F_{y_i'}(x_2). \end{aligned}$$

This criterion for normality is readily established by the same methods as those used by Bliss for the case when E_{12} is an extremal [V, p. 311]. If for a set of multipliers $\lambda_\alpha(x)$ belonging to E_{12} the determinant (2:5) vanishes, then there is a set χ_0, χ_ρ, $c\lambda_\alpha(x)$ having $\chi_0 = 0$ and satisfying the equations (2:6). Hence E_{12} could not be normal. On the other hand if the determinant (2:5) is different from zero for every set of multipliers $\lambda_\alpha(x)$ with which E_{12} satisfies

equations (2:2), then there can be no set χ_0, χ_ρ, $\lambda_\alpha(x)$ with $\chi_0 = 0$ satisfying the equations (2:6). Consequently in this case E_{12} is normal. The last statement in the theorem is readily established by the methods used by Bliss for the case when E_{12} is an extremal [V, p. 311].

THEOREM 2:3. *If an admissible arc E_{12} is normal on $x'x''$ and satisfies the equations (2:2) with a set of multipliers $\lambda_\alpha(x)$, then these multipliers are unique on the interval $x'x''$ except for a constant factor.*

This is a result of the relation (2:3) which implies that the constants $F_{y_i'}(x')$, $F_{y_i'}(x'')$ are unique except for a constant factor since it is possible to select a matrix (2:4) having rank $2n - 1$ on $x'x''$. The multipliers belonging to E_{12} on the interval $x'x''$ are then also unique except for a constant factor since they are completely determined when the set of values $F_{y_i'}(x')$ is specified [IX, p. 680].

3. *The family of extremals.* An *extremal* is an admissible arc with a set of multipliers not vanishing simultaneously,

$$y_i = y_i(x), \qquad \lambda_\alpha = \lambda_\alpha(x) \qquad (x_1 \leqq x \leqq x_2),$$

which have continuous derivatives $y_i'(x)$, $y_i''(x)$, $\lambda_\alpha'(x)$ and satisfy the Euler-Lagrange equations

$$(3:1) \qquad (d/dx)F_{y_i'} - F_{y_i} = 0, \qquad \phi_\alpha = 0.$$

Such an extremal is *non-singular* if the determinant

$$R = \begin{vmatrix} F_{y_i'y_k'} & \phi_{\beta y_i'} \\ \phi_{\alpha y_k'} & 0 \end{vmatrix}$$

is different from zero along it. Along a non-singular extremal E_{12} the equations

$$(3:2) \qquad F_{y_i'}(x, y, y', \lambda) = z_i, \qquad \phi_\alpha(x, y, y') = 0$$

can be solved for the variables y_i', λ_α in a neighborhood of the values (x, y, z) on the arc E_{12}. The solution has the form

$$(3:3) \qquad y_i' = P_i(x, y, z), \qquad \lambda_\alpha = \Lambda_\alpha(x, y, z)$$

and has continuous partial derivatives of the first two orders since the first members of equations (3:2) have such derivatives. The system of equations (3:1) is now equivalent to the system

$$(3:4) \quad dy_i/dx = P_i(x, y, z), \quad dz_i/dx = F_{y_i}[x, y, P(x,y,z), \Lambda(x,y,z)].$$

The functions F, P_i, Λ_α satisfy the homogeneity relations

$$(3:5) \qquad \begin{aligned} F(x, y, y', k\lambda) &= kF(x, y, y', \lambda), \\ P_i(x, y, kz) &= P_i(x, y, z), \\ \Lambda_\alpha(x, y, kz) &= k\Lambda_\alpha(x, y, z) \qquad (k \neq 0). \end{aligned}$$

The first of these relations is a consequence of the definition of F. The last two follow from the fact that the two sets

$$[x,\ y,\ kz,\ P(x,y,z),\ k\Lambda(x,y,z)]\ ,$$
$$[x,\ y,\ kz,\ P(x,y,kz),\ \Lambda(x,y,kz)]$$

satisfy equations (3:2) and must be identical since the solutions P, Λ of these equations are unique when x, y, z are given.

Through every element $(x_0,\ y_0,\ z_0)$ in a neighborhood of the set of values $(x,\ y,\ z)$ on the extremal E_{12} there passes a unique solution

$$(3:6) \qquad y_i = y_i(x,\ x_0,\ y_0,\ z_0), \qquad z_i = z_i(x,\ x_0,\ y_0,\ z_0)$$

of equations (3:4) for which the functions y_i, y_{ix}, z_i, z_{ix} have continuous partial derivatives of the first two orders since the second members of equations (3:4) have such derivatives. The functions $y_i(x,\ x_0,\ y_0,\ z_0)$, $kz_i(x,\ x_0,\ y_0,\ z_0)$ are solutions of equations (3:4), on account of the homogeneity properties (3:5), and have the initial values $(x,\ y,\ z) = (x_0,\ y_0,\ kz_0)$. Since the solutions with these initial values are unique it follows that

$$(3:7) \qquad \begin{aligned} y_i(x,\ x_0,\ y_0,\ kz_0) &= y_i(x,\ x_0,\ y_0,\ z_0), \\ z_i(x,\ x_0,\ y_0,\ kz_0) &= kz_i(x,\ x_0,\ y_0,\ z_0). \end{aligned}$$

Since each curve (3:6) has an initial set at $x = x_{10}$ we lose none of them if we replace x_0 by the fixed value x_{10}. Furthermore, not all the constants z_{i0} are zero at the initial element of E_{12}. We may therefore renumber the solutions (3:6) so that z_{n0} is different from zero. On account of the homo-

geneity relations (3:7) it follows that the initial elements (x_{10}, y_0, z_0), (x_{10}, y_0, kz_0) determine the same curves $y_i = y_i(x, x_{10}, y_0, z_0)$. Hence we lose none of these curves if we assign to z_{no} the fixed value of z_n belonging to E_{1*} at the point 1. Let us for convenience rename the constants $y_{10}, y_{20}, \cdots, y_{no}, z_{10}, \cdots, z_{n-1,0}$ and call them $c_1, c_2, \cdots, c_{2n-1}$ respectively. The family (3:6) then takes the form

$$y_i = y_i(x, c), \qquad z_i = z_i(x, c). \tag{3:8}$$

The equations

$$c_i = y_i(x_{10}, c), \quad c_{n+r} = z_r(x_{10}, c), \quad z_{no} = z_n(x_{10}, c)$$

express the fact that the solutions (3:8) pass through the initial element

$$(x, y_1, \cdots, y_n, z_1, \cdots, z_{n-1}, z_n) = (x_{10}, c_1, \cdots, c_n, c_{n+1}, \cdots, c_{2n-1}, z_{no})$$

and from them we find by differentiation that the determinant

$$\begin{vmatrix} y_{ic_s} & 0 \\ z_{ic_s} & z_i \end{vmatrix} \tag{3:9}$$

takes the value z_{no} at $x = x_{10}$. When we substitute the functions (3:8) in equations (3:3) a set of functions $\lambda_\alpha(x, c)$ is determined, and we have the final result:

THEOREM 3:1. *Every non-singular extremal arc E_{12} is a member of a $(2n - 1)$-parameter family of extremals*

$$(3:10) \qquad y_i = y_i(x, c), \qquad \lambda_\alpha = \lambda_\alpha(x, c) \qquad (x_1 \leqq x \leqq x_2)$$

for special values $(x_1, x_2, c) = (x_{10}, x_{20}, c_0)$. The functions y_i, y_{ix}, z_i, z_{ix}, λ_α have continuous first and second partial derivatives in a neighborhood of the values (x, c) defining E_{12}, and for the special values (x_{10}, c_0) the determinant (3:9) is different from zero.

4. *The second variation for a normal extremal.* Consider a normal extremal arc E_{12} with ends satisfying the conditions $\psi_\rho = 0$. Let ξ_1, ξ_2, $\eta_i(x)$ be a set of admissible variations along E_{12} satisfying the equations $\Psi_\rho(\xi, \eta) = 0$. It can be shown that there is a one-parameter family of admissible arcs

$$(4:1) \qquad y_i = y_i(x, b), \qquad x_1(b) \leqq x \leqq x_2(b)$$

satisfying the end-conditions $\psi_\rho = 0$, containing E_{12} for $b = 0$, and having ξ_1, ξ_2, $\eta_i(x)$ as its variations along E_{12} [IX, p. 695]. The functions $x_1(b)$, $x_2(b)$, $y_i(x, b)$, $y_{ib}(x, b)$ are continuous in a neighborhood of the values (x,b) defining E_{12}, and their derivatives x_{1b}, x_{1bb}, x_{2b}, x_{2bb}, y_{ix}, y_{ixbb}, y_{ibb} have the same property except possibly at the values of x defining the corners of the arc $\eta_i = \eta_i(x)$ $(x_1 \leqq x \leqq x_2)$ in $x\eta$-space.

When the equations

$$
\begin{aligned}
g(b) &= g\left[x_1(b),\ y(x_1(b),\ b),\ x_2(b),\ y(x_2(b),\ b)\right],\\
0 &= \psi_\rho\left[x_1(b),\ y(x_1(b),\ b),\ x_2(b),\ y(x_2(b),\ b)\right],\\
0 &= \phi_\alpha\left[x,\ y(x,\ b),\ y'(x,\ b)\right]
\end{aligned}
$$

are multiplied by constants and multipliers χ_0, χ_ρ, $\lambda_\alpha(x)$, where χ_0, χ_ρ are to be determined later and the functions $\lambda_\alpha(x)$ are the multipliers belonging to E_{12}, it is found by suitable additions that

$$
\begin{aligned}
\chi_0 g(b) &= G\left[x_1(b),\ y(x_1(b),\ b),\ x_2(b),\ y(x_2(b),\ b)\right],\\
0 &= F\left[x,\ y(x,\ b),\ y'(x,\ b),\ \lambda_\alpha(x)\right],
\end{aligned}
$$

where $G = \chi_0 g + \chi_\rho \psi_\rho$. By differentiating these equations for b it follows further that

$$
\begin{aligned}
\chi_0 g'(b) &= (G_{x_1} + y_{i1}' G_{y_{i1}})x_{1b} + G_{y_{i1}} y_{ib}(x_1)\\
&\quad + (G_{x_2} + y_{i2}' G_{y_{i2}})x_{2b} + G_{y_{i2}} y_{ib}(x_2),\\
0 &= F_{y_i} y_{ib} + F_{y_i'} y_{ib}' ,
\end{aligned}
$$

and a second differentiation gives for b = 0

$$
\begin{aligned}
\chi_0 g''(0) &= (G_{x_1} + y_{i1}' G_{y_{i1}})x_{1bb} + G_{y_{i1}} y_{ibb}(x_1)\Big|^{b=0}\\
&\quad + (G_{x_2} + y_{i2}' G_{y_{i2}})x_{2bb} + G_{y_{i2}} y_{ibb}(x_2)\Big|^{b=0} \qquad (4{:}2)\\
&\quad + Q\left[\xi_1,\ \eta(x_1),\ \xi_2,\ \eta(x_2)\right],
\end{aligned}
$$

$$
0 = F_{y_i} y_{ibb} + F_{y_i'} y_{ibb}' \Big|^{b=0} + 2\omega(x,\ \eta,\ \eta'), \qquad (4{:}3)
$$

where Q is a quadratic form in the variations ξ_1, $\eta_i(x_1)$, ξ_2, $\eta_i(x_2)$ of the family (4:1) along E_{12} and

$$
2\omega(x, \eta, \eta') = F_{y_i y_k}\eta_i\eta_k + 2F_{y_i y_k'}\eta_i\eta_k' + F_{y_i' y_k'}\eta_i'\eta_k' . \qquad (4{:}4)
$$

When equation (4:3) is integrated from x_1 to x_2, it is found with the help of the Euler-Lagrange equations (3:1) that

$$(4:5) \qquad 0 = F_{y_i'}y_{ib}b \Big|_{x_1}^{x_2} + \int_{x_1}^{x_2} 2\omega(x, \eta, \eta')dx.$$

From the hypothesis (c) of section 2, and since E_{12} is normal, we can determine the constants χ_0, χ_ρ to satisfy equations (2:6) with $\chi_0 = 1$. Hence by adding equations (4:2) and (4:5) it follows that the second variation I_2 along E_{12} can be expressed in the form

$$(4:6) \qquad \begin{aligned} I_2 = g''(0) &= Q[\xi_1, \eta(x_1), \xi_2, \eta(x_2)] \\ &+ \int_{x_1}^{x_2} 2\omega(x, \eta, \eta')dx, \end{aligned}$$

and this expression must be $\geqq 0$ for every set of admissible variations ξ_1, ξ_2, $\eta_i(x)$ along E_{12} satisfying the conditions $\Psi_\rho(\xi, \eta) = 0$.

Since E_{12} is normal the relation (2:3) and Theorem 2:2 imply that every set of admissible variations ξ_1, ξ_2, $\eta_i(x)$ along E_{12} satisfying the conditions $\Psi_\rho = 0$ also satisfies the equations $\xi_1 = \eta_i(x_1) = \xi_2 = \eta_i(x_2) = 0$. Hence in the expression (4:6) the value of the quadratic form Q is always zero, and we have the following theorem:

THEOREM 4:1. _Along a normal extremal arc E_{12} with ends satisfying the conditions $\psi_\rho = 0$ the second variation is always expressible in the form_

$$I_2 = \int_{x_1}^{x_2} 2\omega(x, \eta, \eta')dx$$

for all admissible variations ξ_1, ξ_2, $\eta_i(x)$ _satisfying_

<u>the equations</u> $\Psi_\rho = 0$, <u>where</u> $2\,\omega$ <u>is the quadratic form</u> (4:4). <u>If</u> $g(E_{12})$ <u>is to be a minimum for the problem of Mayer as here proposed, then this second variation must be</u> $\geqq 0$ <u>for every set of admissible variations</u> $\eta_i(x)$ <u>satisfying the relations</u>

$$(4:7) \qquad \eta_i(x_1) = \eta_i(x_2) = 0.$$

Since the functions $\eta_i(x)$ satisfy the differential equations of variation

$$(4:8) \qquad \Phi_\alpha(x, \eta, \eta') = \phi_{\alpha y_i}\eta_i + \phi_{\alpha y_i'}\eta_i' = 0$$

it is clear that the properties of the second variation suggest a minimum problem which is a problem of Lagrange [cf. VI, p. 16] namely, that of minimizing I_2 in the class of arcs

$$(4:9) \qquad \eta_i = \eta_i(x) \qquad (x_1 \leqq x \leqq x_2)$$

satisfying equations (4:8) and passing through the fixed points $(x_1, 0)$, $(x_2, 0)$ in $x\eta$-space as indicated by equations (4:7). One readily verifies that this problem is abnormal since, as was seen in Section 2, the rank of the matrix (2:4) cannot exceed $2n - 1$ on E_{12}. However, by a suitable modification of the end-conditions the problem can be made normal. For this purpose we replace the condition that the arc (4:9) passes through the fixed points (x_1, o), $(x_2, 0)$ in $x\eta$-space by the conditions

$$(4:10) \qquad x_1 - \alpha_1 = \eta_i(x_1) = x_2 - \alpha_2 = \eta_\chi(x_2) = 0 \qquad (\chi = 1, \cdots, n;\ \chi \neq p),$$

where p is chosen so that $F_{y_p'}(x_2) \neq 0$. The two sets of end-conditions are equivalent since the relation (2:3) implies that $\eta_p(x_2) = 0$ whenever the conditions (4:10) are satisfied.

To prove that the new accessory problem just described is normal we use the fact that since E_{12} is normal there is a determinant of the form $|\Psi_\rho(\xi^\sigma, \eta^\sigma)|$ which is different from zero on E_{12}. The matrix of this determinant is the product of two matrices, the first of which is formed by deleting the first row of the matrix (2:5) and has rank 2n + 1, and the second of which is a matrix having 2n + 1 columns of the form

$$\xi_1^\sigma, \eta_i^\sigma(x_1), \; \xi_2^\sigma, \eta_i^\sigma(x_2). \tag{4:11}$$

This second matrix must also have rank 2n + 1 if the original determinant is to be different from zero, and the determinant formed from this second matrix by leaving out the row of elements $\eta_p^\sigma(x_2)$ must be different from zero, as one readily sees with the help of the relation (2:3). This last determinant is however one of the form whose non-vanishing insures the normality of the accessory problem with end-conditions (4:10).

The Euler-Lagrange equations for the $x\eta$-problem are the equations

$$(d/dx)\Omega_{\eta_i'} - \Omega_{\eta_i} = 0, \quad \Phi_\alpha(x, \eta, \eta') = 0, \tag{4:12}$$

where $\Omega\,(x, \eta, \eta', \mu) = \mu_0 \omega + \mu_\alpha \Phi_\alpha$. These equations are known as the <u>accessory</u> <u>equations</u> for the original Mayer problem.

THEOREM 4:2. <u>If the functions $\mu_0 = 1$, $\mu_\alpha(x)$ are a set of multipliers with which an admissible arc (4:9) for the $x\eta$-problem satisfies equations (4:12), then every set of functions $\rho_0 = 1$, $\rho_\alpha(x)$ having this property is of the form $\rho_0 = 1$, $\rho_\alpha(x) = \mu_\alpha(x) + k\lambda_\alpha(x)$, where the functions $\lambda_\alpha(x)$ are the multipliers for E_{12} and k is an arbitrary constant.</u>

This follows because if $\rho_0 = 1$, $\rho_\alpha(x)$ are a second set of multipliers for the arc (4:9), then the differences $\rho_\alpha(x) - \mu_\alpha(x)$ must be multipliers for the original problem and hence be of the form $\rho_\alpha(x) - \mu_\alpha(x) = k\,\lambda_\alpha(x)$, since E_{12} is normal. This proves the theorem [cf. VI, p. 19].

An admissible arc (4:9) having associated with it a set of multipliers μ_0, $\mu_\alpha(x)$ with which it satisfies equations (4:12) will also satisfy the <u>transversality</u> <u>condition</u> for the accessory problem just described if it satisfies the relation $\Omega_{\eta_p'}(x_2) = 0$ [IX, p. 693]. Since E_{12} is normal and $F_{y_p'}(x_2) \neq 0$ it follows that a solution $\eta_i(x)$, $\rho_0 = 1$, $\rho_\alpha = \mu_\alpha(x) + k\,\lambda_\alpha(x)$ of equations (4:12) satisfies the transversality condition $\Omega_{\eta_p'}(x_2) = 0$ for a suitably selected value of the constant k.

Let us now assume that E_{12} is also non-singular. Then the determinant R is different from zero along E_{12}, and the

equations

$$\Omega_{\eta_i'}(x, \eta, \eta', \mu) = \zeta_i, \quad \Phi_\alpha(x, \eta, \eta') = 0$$

with $\mu_0 = 1$ can be solved for the variables η_i', μ_α. The solution has the form

$$\eta_i' = H_i(x, \eta, \zeta), \qquad \mu_\alpha = M_\alpha(x, \eta, \zeta),$$

and the accessory equations (4:12) with $\mu_0 = 1$ are now equivalent to the equations

$$(4:13) \qquad \begin{aligned} d\eta_i/dx &= H_i(x, \eta, \zeta), \\ d\zeta_i/dx &= \Omega_{\eta_i}[x, \eta, H(x, \eta, \zeta), M(x, \eta, \zeta)] \end{aligned}$$

which are linear and homogeneous in the variables η_i, ζ_i. They have the solution $\eta_i \equiv 0$, $\zeta_i = z_i(x)$, where $z_i(x)$ are the values of the derivatives $F_{y_i'}$ along E_{12}, since the corresponding values $\eta_i \equiv 0$, $\mu_\alpha = \lambda_\alpha$ reduce the first equations (4:12) to the Euler-Lagrange equations (3:1). It is known that for equations (4:13) a set of 2n - 1 solutions u_{is}, v_{is}, whose determinant

$$(4:14) \qquad \begin{vmatrix} u_{is} & 0 \\ v_{is} & z_i \end{vmatrix}$$

is different from zero for one value of x, has that determinant different from zero for all values of x. Furthermore every solution (η_i , ζ_i) of equations (4:13) is expressible in the form

$$(4{:}15)\qquad \eta_i = c_s u_{is}, \qquad \zeta_i = c_s v_{is} + kz_i,$$

where c_s, k are constants [IV, pp. 153 - 4]. One readily verifies that the columns of the determinant (3:9) are a set of solutions of equations (4:13) like those in the columns of (4:14) [IX, p. 726].

As an immediate consequence of the relation (4:15) it follows that there is one and only one solution (η_i, ζ_i) of equations (4:12) taking prescribed values η_{io}, ζ_{io} at a given value x_0. In particular the only solution taking the values $\eta_{io} = \zeta_{io} = 0$ at $x = x_0$ is the solution $\eta_i \equiv \zeta_i \equiv 0$. Furthermore, since E_{12} is normal the only solution having $\eta_i \equiv 0$ on x_1x_2 is the solution $\eta_i \equiv 0$, $\zeta_i = kz_i(x)$. The same is true on a sub-interval $x'x''$ provided E_{12} is normal on this sub-interval.

5. <u>The necessary condition of Mayer</u>. A value $x_3 \neq x_1$ is said to define a point 3 <u>conjugate</u> to 1 on E_{12} if there exists a solution $\eta_i = u_i(x)$, $\mu_0 = 1$, $\mu_\alpha = \rho_\alpha(x)$ of equations (4:12) whose functions $u_i(x)$ satisfy the relations $u_i(x_1) = u_i(x_3) = 0$ but are not all identically zero on x_1x_3.

IV. THE NECESSARY CONDITION OF MAYER. <u>If E_{12} is a normal non-singular minimizing extremal arc then between the points 1 and 2 on E_{12} there can be no point 3 conjugate to 1 defined by a value x_3 such that E_{12} is normal on the interval x_3x_2.</u>

If there were a solution $\eta_i = u_i(x)$, $\mu_0 = 1$, $\mu_\alpha = \rho_\alpha(x)$ of equations (4:12) whose functions $u_i(x)$ vanish at

x_1 and x_3 but are not all identically zero on x_1x_3, then for the functions $\eta_i(x)$, μ_0, $\mu_\alpha(x)$ defined by the equations

$$
\begin{aligned}
&\eta_i(x) \equiv u_i(x), \quad \mu_0 = 1, \quad \mu_\alpha(x) \equiv \rho_\alpha(x) \quad \text{on } x_1x_3, \\
&\eta_i(x) \equiv 0 \quad\quad\;\; , \quad \mu_0 = 1, \quad \mu_\alpha(x) \equiv 0 \quad\quad\; \text{on } x_3x_2
\end{aligned}
\tag{5:1}
$$

the second variation I_2 would take the value zero [IX, p. 726]. It follows that the arc

$$
\eta_i = \eta_i(x) \qquad (x_1 \leqq x \leqq x_2)
\tag{5:2}
$$

would be a minimizing arc for the $x\eta$-problem since E_{12} is to be a solution of the original problem. Hence there would be associated with the arc (5:2) a function $\Omega = \omega + \mu_\alpha \Phi_\alpha$ with which it would satisfy the accessory equations (4:12), the transversality condition $\Omega_{\eta'_p}(x_2) = 0$, and the condition that the derivatives $\Omega_{\eta'_i}(x)$ are continuous on the interval x_1x_2. If E_{12} is normal on the interval x_3x_2 then, as was seen above, the most general multipliers possible for the functions $\eta_i(x)$ would have the forms $\mu_0 = 1$, $\mu_\alpha = d\,\lambda_\alpha(x)$ on the interval x_3x_2, and on account of the transversality condition $\Omega_{\eta'_p}(x_2) = 0$ the constant d would be zero since $F_{y'_p}(x_2) \neq 0$. Hence at $x = x_3$ the corner condition would require

$$
\Omega_{\eta_i'}(x_3 - 0) = \omega_{\eta_i'}(x, u, u') + \mu_\alpha \phi_{\alpha y_i'} \Big|^{x_3} = 0.
$$

It follows that there would exist for the arc (5:2) a set of multipliers $\mu_0 = 1$, $\mu_\alpha(x)$ such that at $x = x_3$ the functions $\zeta_i = \Omega_{\eta_i'}(x, u, u', \mu)$ vanish as well as $\eta_i = u_i$.

cally which is not the case, and the theorem is therefore established [cf. VI, p. 18].

6. The determination of conjugate points. Consider a non-singular, normal extremal arc E_{12} that is normal on every subinterval x_1x_3.

THEOREM 6:1. *Let u_{is}, v_{is} be $2n - 1$ solutions of equations (4:13) whose determinant (4:14) is different from zero at $x = x_1$. A value $x_3 \neq x_1$ determines a point 3 conjugate to 1 on E_{12} if and only if the matrix*

$$\left\| \begin{matrix} u_{is}(x_3) \\ u_{is}(x_1) \end{matrix} \right\| \tag{6:1}$$

has rank $< 2n - 1$.

This theorem is a simple extension of a theorem given by Larew and can be proved by the same methods [VI, p. 20].

If now we select $2n - 1$ solutions u_{is}, v_{is} of equations (4:13), as in Theorem 6:1, and such that at $x = x_1$ the functions $u_{is}(x)$ have the values

$$u_{ir}(x_1) = 0, \quad u_{i,n-1+k}(x_1) = \delta_{ik} \quad (\delta_{ii} = 1,\ \delta_{ik} = 0 \text{ for } i \neq k),$$

then it is clear that the matrix (6:1) for this set has rank $2n - 1$ if and only if the matrix $\| u_{ir}(x_3) \|$ has rank $n - 1$. With this in mind we can prove the following theorem:

THEOREM 6:2. *Let u_{ik}, v_{ik} be n solutions of equations (4:13) which at $x = x_1$ satisfy the relations*

$$u_{ir}(x_1) = 0, \quad |v_{ir}(x_1) \quad z_i(x_1)| \neq 0, \quad u_{in}(x_1) = z_i(x_1),$$
$$v_{in}(x_1) = 0.$$

<u>A value</u> $x_3 \neq x_1$ <u>determines a point</u> 3 <u>conjugate to</u> 1 <u>on</u> E_{12} <u>if and only if</u> $D(x_3) = 0$, <u>where</u> $D(x) = |u_{ik}(x)|$.

The theorem follows at once from our previous considerations if we show that $D(x_3)$ vanishes if and only if the matrix $\| u_{ir}(x_3) \|$ has rank $< n - 1$. If now $D(x_3) = 0$, then there exist constants a_k, not all zero, such that $u_{ik}(x_3)a_k = 0$. On account of the relation (2:3) for the functions $\eta_i(x) = u_{ik}(x)a_k$ and the values of u_{ik} at $x = x_1$ it follows that

$$0 = z_i(x_3) \, u_{ik}(x_3)a_k = z_i(x_1) \, u_{ik}(x_1)a_k = z_i(x_1) \, z_i(x_1)a_n.$$

Hence $a_n = 0$, and the matrix $\| u_{ir}(x_3) \|$ has rank $< n - 1$. The converse is immediate, and the theorem is established.

7. <u>Mayer fields and a fundamental sufficiency theorem</u>. The importance of the introduction of the notion of an $(n + 1)$-dimensional field in the space of points $(x, y_1, \cdots, y_n)$ for the problems of Mayer will be seen from the following considerations.

DEFINITION OF A MAYER FIELD. A <u>Mayer field</u> for the problem considered in this paper is a region $\mathfrak{F}$ in xy-space containing only interior points and having associated with it a set of functions $p_i(x, y)$, $\lambda_\alpha(x, y)$ with the following properties:

(a) they have continuous first partial derivatives in $\mathfrak{F}$;

(b) the sets $[x, y, p(x,y)]$ defined by the points (x, y) in $\mathfrak{F}$ are all admissible;

(c) the integral

$$I^* = \int \left\{F(x,y,p,\lambda)dx + (dy_i - p_i dx)F_{y_i'}(x,y,p,\lambda)\right\}$$

formed with these functions is independent of the path in $\mathfrak{F}$.

This definition of a field is precisely the one given by Bliss for the problem of Lagrange except for the form of the function F [IX, p. 730]. It should be noted that for the problem of Mayer here discussed the function $F(x, y, p, \lambda)$ vanishes identically in $\mathfrak{F}$, which is not in general true for the problems of Lagrange. Bliss has shown that the solutions $y_i(x)$ of the equations $dy_i/dx = p_i(x, y)$ are extremals with multipliers $\lambda_\alpha(x, y(x))$, called extremals of the field. It is clear that the value of I* is zero along every extremal of the field.

THEOREM 7:1. *If E_{12} is a normal extremal arc of a field $\mathfrak{F}$ with ends satisfying the conditions $\psi_\rho = 0$, then there is a neighborhood N of the ends of E_{12} in $(x_1y_1x_2y_2)$-space such that for every admissible arc C_{34} in $\mathfrak{F}$ with ends in N satisfying the conditions $\psi_\rho = 0$ the formula*

$$g(C_{34}) - g(E_{12}) = (1/\lambda_0)\int_{x_3}^{x_4} E[x,y,p(x,y),\lambda(x,y), y']dx \tag{7:1}$$

holds, where λ_0 is a suitably chosen positive constant,

$$E(x,y,p,\lambda,y') = F(x,y,y',\lambda) - F(x,y,p,\lambda) - (y_i' - p_i)\, F_{y_i'}(x,y,p,\lambda),$$

and the arguments $y_i(x)$, $y_i'(x)$ occurring in the integrand are those belonging to C_{34}.

As a first step in the proof consider the equations

$$g(x_1, y_1, x_2, y_2) = g, \quad \psi_\rho(x_1, y_1, x_2, y_2) = 0.$$

By hypothesis they are satisfied by the set $[x_1, y_1, x_2, y_2, g(E_{12})]$ belonging to E_{12}. Since the determinant (2:1) is different from zero these equations have solutions of the form

$$(7:2) \quad x_1 = x_1(g), \quad y_{i1} = y_{i1}(g), \quad x_2 = x_2(g), \quad y_{i2} = y_{i2}(g)$$

which have continuous second derivatives in a neighborhood of the value $g = g(E_{12})$. Furthermore, in a sufficiently small neighborhood N of the ends of E_{12} the only solutions are those defined by equations (7:2). These equations define two arcs A, B through the ends of E_{12}.

The equations

$$\begin{aligned}
\chi_0 g_{x_1} + \chi_\rho \psi_{\rho x_1} &= - p_i F_{y_i'}(x, y, p, \lambda)\Big|^1, \\
\chi_0 g_{y_{i1}} + \chi_\rho \psi_{\rho y_{i1}} &= F_{y_i'}(x, y, p, \lambda)\Big|^1, \\
\chi_0 g_{x_2} + \chi_\rho \psi_{\rho x_2} &= p_i F_{y_i'}(x, y, p, \lambda)\Big|^2,
\end{aligned}$$

$$\chi_0 g_{y_{12}} + \chi_\rho \psi_{\rho y_{i2}} = - F_{y_i'}(x, y, p, \lambda) \Big|^2 ,$$

where the variables x_1, y_{i1}, x_2, y_{i2} are replaced by the right members of equations (7:2), determine continuous functions $\chi_0(g)$, $\chi_\rho(g)$. When they are multiplied by the differentials dx_1, dy_{i1}, dx_2, dy_{i2} belonging to the arcs A, B and added it is found that

$$\chi_0 \, dg = -F_{y_i'}(dy_i - p_i dx) \Big|_1^2 . \tag{7:3}$$

In order to compare the values of g for the arcs E_{12} and C_{34} this last equation may be integrated from $g = g(E_{12})$ to $g = g(C_{34})$. By then applying the first law of the mean to the left member an equation of the form

$$\lambda_0 \left[g(C_{34}) - g(E_{12}) \right] = I^*(A_{13}) - I^*(B_{24}) \tag{7:4}$$

is obtained, where λ_0 is a suitably selected mean value of the function $\chi_0(g)$ on E_{12}. Since E_{12} is normal we may suppose $\chi_0 = 1$ on E_{12}, according to the agreement made in Section 2. Consequently the neighborhood N can be chosen so small that $\chi_0(g) > 0$ and hence $\lambda_0 > 0$ in N. Furthermore, since I* is independent of the path in $\mathfrak{F}$ it is clear that

$$I^*(A_{13}) - I^*(B_{24}) = I^*(E_{12}) - I^*(C_{34}) = - I^*(C_{34}), \tag{7:5}$$

the last equality being valid since I* vanishes identically along the extremal E_{12} of the field. The theorem now follows at once from equations (7:4) and (7:5) since, as is easily seen, the value of $- I^*(C_{34})/\lambda_0$ is equal to the value of the

second member of equation (7:1).

It is now possible to prove the following important theorem:

THEOREM 7:2. A FUNDAMENTAL SUFFICIENCY THEOREM. Let a normal extremal arc E_{12} be an extremal of a field $\mathfrak{F}$. Suppose that the ends of E_{12} satisfy the conditions $\psi_\rho = 0$ and that there is a neighborhood N of those ends in $(x_1 y_{i1} x_2 y_{i2})$-space such that no other extremal of the field has ends in N satisfying the equations $\psi_\rho = 0$. If at each point of $\mathfrak{F}$ the condition

$$E\,[x,\ y,\ p(x,y),\ \lambda\,(x,y),\ y']>0$$

holds for every admissible set $(x, y, y') \neq (x, y, p)$, then the neighborhood N can be so restricted that the inequality $g(C_{34}) > g(E_{12})$ is true for every admissible arc C_{34} in $\mathfrak{F}$ with ends in N satisfying the conditions $\psi_\rho = 0$ and not identical with E_{12}.

To prove this restrict N so as to be effective as in Theorem 7:1. It follows at once from Theorem 7:1 that the inequality $g(C_{34}) \geqq g(E_{12})$ is necessarily satisfied by every admissible arc C_{34} in $\mathfrak{F}$ with ends in N satisfying the conditions $\psi_\rho = 0$. The equality sign is appropriate only when the E-function vanishes along C_{34}, that is, only when $y_i' = p_i$ at each point of C_{34}. But in that case C_{34} would be an extremal of the field and would coincide with E_{12} since E_{12} is the only extremal of the field with ends in N satisfying the conditions $\psi_\rho = 0$.

8. <u>An auxiliary theorem.</u> A normal extremal arc E_{12} is said to satisfy the <u>Clebsch condition III'</u> if at each element (x, y, y', λ) on it the inequality

$$F_{y_i'y_k'}\pi_i\pi_k > 0$$

holds for every set $(\pi_1, \cdots, \pi_n) \neq (0, \cdots, 0)$ which is a solution of the equations $\phi_{\alpha y_i'}\pi_i = 0$. The arc E_{12} satisfies the <u>Mayer condition IV'</u> if there is no point 3 conjugate to 1 on E_{12} between 1 and 2 or at 2.

In this section we propose to construct n solutions U_{ik}, V_{ik} of equations (4:13) whose determinant $|U_{ik}(x)|$ is as stated in Theorem 8:1 below different from zero on x_1x_2. To do this we consider a normal extremal arc E_{12} that is normal on every sub-interval x_1x_3 and satisfies the conditions III', IV' just described. From the condition III' we conclude that E_{12} is non-singular [IX, p. 735].

LEMMA 8:1. <u>There is an interval</u> $x_1 < x \leqq x_1 + h$ <u>on which there is no point</u> 3 <u>conjugate to</u> 1 <u>on</u> E_{12}.

This lemma is readily proved by the methods used by Bliss to establish the corresponding theorem for the problem of Lagrange [IX, pp. 737 - 740]. Bliss makes the stronger assumption that E_{12} is normal on every sub-interval $x'x''$, a restriction which is useful if we wish to show that there are no pairs of conjugate points whatsoever on E_{12} defined by values $x'x''$ on an interval $x_1 \leqq x \leqq x_1 + h$. It can, however, be replaced by the weaker hypothesis that E_{12} is

normal on every sub-interval x_1x_3 if we wish to consider only the points 3 conjugate to 1 on E_{12}.

For every pair of solutions (η_i, ζ_i), (u_i, v_i) of equations (4:13) it is known that the expression $\eta_i v_i - u_i \zeta_i$ is a constant. If this constant is zero, then the two solutions are called <u>conjugate solutions</u> of equations (4:13). A set of n mutually conjugate solutions of equations (4:13) is said to form a <u>conjugate system</u> of solutions.

Consider now the system of solutions u_{ik}, v_{ik} of equations (4:13) defined in Theorem 6:2. One readily verifies that this system forms a conjugate system if the functions $v_{ik}(x)$ are modified so that they satisfy the relation $z_i(x_1)\cdot v_{ik}(x_1) = 0$. This can be done by adding to the solution u_{ik}, v_{ik} suitable multiples of the solution $\eta_i \equiv 0$, $\zeta_i = z_i(x)$. Furthermore, since E_{12} satisfies the condition IV' it follows from Theorem 6:2 and Lemma 8:1 that the determinant $|u_{ik}(x)|$ is different from zero on the interval $x_1 < x \leqq x_2$. When the matrices $\| u_{ik} \|$, $\| v_{ik} \|$ are multiplied on the right by the inverse of the matrix $\| u_{ik}(x_2) \|$ a new conjugate system η_{ik}, ζ_{ik} is formed which takes the values δ_{ik}, B_{ik} at $x = x_2$, where δ_{ik} equals 0 or 1 according as $i \neq k$ or $i = k$, and $B_{ik} = B_{ki}$. It is clear that the determinant $|\eta_{ik}(x)|$ is also different from zero on the interval $x_1 < x \leqq x_2$. Hence the n-parameter family of solutions of equations (4:13)

$$\eta_i = \eta_{ik}a_k, \qquad \zeta_i = \zeta_{ik}a_k \qquad (x_1 \leqq x \leqq x_2) \tag{8:1}$$

simply covers a region $\mathfrak{F}$ of points $(x, \eta_1, \cdots, \eta_n)$ whose x-coordinates lie on the interval $x_1 < x \leqq x_2$. Each arc of this family intersects the hyperplane $x = x_2$ in points whose η-coordinates are the parameters a_k defining the arc. Furthermore, on the hyperplane $x = x_2$ the Hilbert integral $I_2{*}$ for the $x\eta$-problem defined by the family (8:1) takes the form

$$I_2{*} = \int 2\,\zeta_i d\eta_i = \int 2\,B_{ik}a_k da_i = \int d(B_{ik}a_i a_k)$$

and hence is independent of the path. It follows that the family (8:1) defines a field $\mathfrak{F}$ [IX, p. 733], and the following lemma is established:

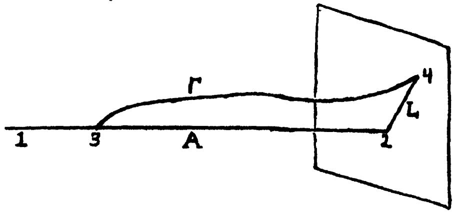
LEMMA 8:2. *If η_{ik}, ζ_{ik} is a conjugate system of solutions taking at $x = x_2$ the values δ_{ik}, B_{ik} just defined, then the determinant $|\eta_{ik}(x)|$ is different from zero on the interval $x_1 < x \leqq x_2$. Furthermore the n-parameter family (8:1) of solutions of the accessory equations defines a Mayer field over a region $\mathfrak{F}$ of points $(x, \eta_1, \cdots, \eta_n)$ whose x-coordinates lie on the interval $x_1 < x \leqq x_2$.*

LEMMA 8:3. *For every extremal Γ_{34} for the $x\eta$-problem joining points $(x, \eta) = (x_3, 0)$ and $(x, \eta) = (x_2, a)$, with $x_1 \leqq x_3 < x_2$, the relation*

$$I_2(\Gamma_{34}) - B_{ik}a_i a_k \geqq 0 \tag{8:2}$$

holds where

$$I_2 = \int 2\,\omega(x, \eta, \eta')dx.$$



Consider first the case when $x_3 > x_1$. According to Lemma 8:2 the Hilbert integral I_8* for the integral I_8 is independent of the path in $\mathfrak{F}$. Hence

$$(8:3)\qquad \begin{aligned} I_8*(\Gamma_{34}) &= I_8*(A_{32}) + I_8*(L_{24}) \\ &= \int_{L_{24}} 2B_{ik}\,\eta_i d\eta_k = B_{ik}a_i a_k . \end{aligned}$$

Since Γ_{34} is admissible it follows that

$$(8:4)\qquad I_8(\Gamma_{34}) - I_8*(\Gamma_{34}) = \int_{\Gamma_{34}} E_\Omega\, dx,$$

where E_Ω is the Weierstrass E-function formed for the function 2Ω. By the use of Taylor's expansion one readily verifies that the condition III' on E_{18} implies that $E_\Omega \geqq 0$ along Γ_{34}. Hence from equations (8:3) and (8:4) it is clear that the inequality (8:2) is true whenever $x_3 > x_1$. If now $x_3 = x_1$ then Γ_{34} is an extremal of the field and by direct integration it is found that $I_8(\Gamma_{34}) = B_{ik}a_i a_k$. Hence the lemma is established.

The following theorem gives us the result described at the beginning of this section.

THEOREM 8:1. <u>Let</u> U_{ik}, V_{ik} <u>be a conjugate system of solutions of equations (4:13) having at</u> $x = x_8$ <u>the initial</u>

values δ_{ik}, $H_{ik} = B_{ik} - \delta_{ik}$, *where* δ_{ik}, B_{ik} *are the values described above. For such a system the determinant* $|U_{ik}(x)|$ *is different from zero on the whole interval* $x_1 \leqq x \leqq x_2$ *and* $H_{ik} = H_{ki}$.

In the first place $|U_{ik}(x_2)| = 1$. If now $|U_{ik}(x)|$ vanishes for a value x_3 ($x_1 \leqq x_3 < x_2$), then there exist constants a_k, not all zero, such that $U_{ik}(x_3)a_k = 0$. The equations

$$\eta_i = U_{ik}a_k\,, \qquad \zeta_i = V_{ik}a_k$$

define an arc Γ_{34} as in Lemma 8:3. By direct integration it is found that for this arc

$$I_2(\Gamma_{34}) - B_{ik}a_ia_k = (B_{ik} - \delta_{ik})a_ia_k - B_{ik}a_ia_k = -a_ia_i < 0.$$

This contradicts the result obtained in Lemma 8:3. Hence $|U_{ik}(x_3)|$ is different from zero on the whole interval x_1x_2 as was to be proved.

9. *The construction of a field.* In order to construct a field we need the following theorem:

THEOREM 9:1. *Suppose that an n-parameter family of extremals*

$$(9:1) \qquad y_i = y_i(x, a_1, \cdots, a_n), \quad \lambda_\alpha = \lambda_\alpha(x, a_1, \cdots, a_n)$$

is intersected by an n-dimensional manifold

$$(9:2) \qquad x = x_1(a_1, \cdots, a_n), \quad y_i = y_i\left[x_1(a_1,\cdots,a_n), a_1,\cdots,a_n\right]$$

and simply covers a region $\mathfrak{F}$ *of xy-space containing only*

interior points. If the parameter values of the extremal through the point (x, y) are denoted by $a_i(x, y)$, then the region $\mathfrak{F}$ is a field with slope-functions and multipliers

$$p_i(x, y) = y_{ix}[x, a(x,y)], \quad \lambda_\alpha(x,y) = \lambda_\alpha[x, a(x,y)] \tag{9:3}$$

provided that the integral I^* is independent of the path on the n-dimensional manifold (9:2).

This theorem has been established by Bliss for the problem of Lagrange [IX, p. 733]. The proof is the same for the problem considered here.

THEOREM 9:2. *If a normal extremal arc E_{12} is normal on every sub-interval x_1x_3 and satisfies the conditions III', IV', then E_{12} is a member of an n-parameter family of extremals (9:1) whose determinant $|y_{ia_k}|$ is different from zero along E_{12}. Furthermore E_{12} is an extremal arc of a field $\mathfrak{F}$ simply covered by the family.*

To prove this let $W(a_1, \cdots, a_n)$ be a function of the form

$$W(a) = z_{12}a_1 + (1/2)H_{ik}(a_i - y_{i2})(a_k - y_{k2}), \tag{9:4}$$

where the constants y_{i2}, z_{12} are the values of the functions $y_i(x)$, $z_1(x)$ defining E_{12} at $x = x_2$, and the H_{ik} are the numbers belonging to the conjugate system U_{ik}, V_{ik} defined in Theorem 8:1. When in equations (3:6) the set (x_0, y_{io}, z_{io}) is replaced by the set (x_2, a_i, W_{a_i}), an n-parameter family of extremals

$$(9:5)\qquad \begin{aligned} y_i &= y_i(x,\ x_2,\ a,\ W_a) = y_i(x,\ a), \\ z_i &= z_i(x,\ x_2,\ a,\ W_a) = z_i(x,\ a) \end{aligned}$$

is defined and contains E_{12} for the special values $a_i = y_{i2}$. The multipliers $\lambda_\alpha(x, a)$ associated with this family are determined by equations (3:3). Furthermore, since each extremal (9:5) defined by parameter values a_i has on it the element (x_2, a_i, W_{a_i}), it follows that $y_{ia_k} = \delta_{ik}$, $z_{ia_k} = W_{a_i a_k} = H_{ik}$ at $x = x_2$. Hence from Theorem 8:1 we conclude that the determinant $|y_{ia_k}|$ is different from zero along each extremal of the family (9:5). This family, therefore, simply covers a neighborhood $\mathfrak{F}$ of E_{12}. Moreover, on the hyperplane $x = x_2$ the Hilbert integral I^* can be expressed in the form

$$I^* = \int F_{y_i'} dy_i = \int W_{a_i} da_i = \int dW$$

and hence is independent of the path. Theorem 9:1 now justifies the theorem that was to be proved.

THEOREM 9:3. *Let a normal extremal arc E_{12} be a member of an n-parameter family of extremals (9:1) whose determinant $|y_{ia_k}|$ is different from zero along E_{12}. If the ends of E_{12} satisfy the conditions $\psi_\rho = 0$, then there is a neighborhood N of these ends in $(x_1 y_{i1} x_2 y_{i2})$-space such that E_{12} is the only extremal of the family with ends in N satisfying the conditions $\psi_\rho = 0$.*

To prove this let E_{12} be a member of the family (9:1) for the special parameter values (x_{10}, x_{20}, a_0). By hypothesis

these values satisfy the equations

$$\psi_\rho(x_1,\ x_2,\ a) = \psi_\rho\left[x_1,\ y(x_1,a),\ x_2,\ y(x_2,a)\right] = 0.$$

The theorem now follows at once from implicit function theorems if we can show that the matrix

$$(9:6)\quad \left\| \psi_{\rho x_1} + y_{i1}'\psi_{\rho y_{i1}} \qquad \psi_{\rho x_2} + y_{i2}'\psi_{\rho y_{i2}} \qquad \psi_{\rho y_{i1}} y_{ia_k}(x_1) + \psi_{\rho y_{i2}} y_{ia_k}(x_2) \right\|$$

has rank $n + 2$ on E_{12}. To do this suppose that it had rank less than $n + 2$. Then there would exist constants b_1, b_2, c_k, not all zero such that the relations

$$\begin{aligned}(\psi_{\rho x_1} + y_{i1}'\psi_{\rho y_{i1}})b_1 + (\psi_{\rho x_2} + y_{i2}'\psi_{\rho y_{i2}})b_2 + \psi_{\rho y_{i1}} y_{ia_k}(x_1)c_k \\ + \psi_{\rho y_{i2}} y_{ia_k}(x_2)c_k = 0, \\ F_{y_i'}(x_1)y_{ia_k}(x_1)c_k - F_{y_i'}(x_2)y_{ia_k}(x_2)c_k = 0\end{aligned}$$

would hold on E_{12}. The last equation is precisely the relation (2:3) for the admissible variations $\eta_i = y_{ia_k}c_k$. On account of the normality of E_{12} the determinant (2:5) is different from zero on E_{12}. Hence we would have

$$b_1 = b_2 = y_{ia_k}(x_{10},\ a_0)c_k = y_{ia_k}(x_{20},\ a_0)c_k = 0.$$

But this is impossible since the determinant $|y_{ia_k}|$ is different from zero along E_{12}. The matrix (9:6) therefore has rank $n + 2$ on E_{12}, and the theorem is established.

10. Sufficient conditions for relative minima. The condition I is defined in Section 2; the Clebsch condition III' and the Mayer condition IV' in Section 8. A normal

minimizing arc E_{12} is said to satisfy the Weierstrass condition $II_{\mathfrak{N}}'$ if at each element (x, y, y', λ) in a neighborhood $\mathfrak{N}$ of those belonging to E_{12} the inequality

$$E(x, y, y', \lambda, Y') > 0$$

holds for every admissible element $(x, y, Y') \neq (x, y, y')$.

THEOREM 10:1. SUFFICIENT CONDITIONS FOR A STRONG RELATIVE MINIMUM. Let E_{12} be an admissible arc without corners and with ends satisfying the conditions $\psi_\rho = 0$. If E_{12} is normal relative to the end-conditions $\psi_\rho = 0$, is normal on every sub-interval x_1x_3 of x_1x_2, and satisfies the conditions I, $II_{\mathfrak{N}}'$, III', IV', then there are neighborhoods $\mathfrak{F}$ of E_{12} in xy-space and N of the ends of E_{12} in $(x_1y_1x_2y_2)$-space such that the inequality $g(C_{34}) > g(E_{12})$ holds for every admissible arc C_{34} in $\mathfrak{F}$ with ends in N satisfying the conditions $\psi_\rho = 0$ and not identical with E_{12}.

To prove this theorem we first notice that the condition I and the normality of E_{12} imply a unique set of multipliers $\lambda_\alpha(x)$ and constants c_i with which E_{12} satisfies equation (2:2) and for which $\chi_0 = 1$, as agreed upon in Theorem 2:2. The condition III' implies further that E_{12} is non-singular and hence must be a single extremal arc since it has no corners [IX, p. 735]. According to Theorem 9:2 we now see that E_{12} is an extremal of a field $\mathfrak{F}$ with slope functions and multipliers $p_i(x, y)$, $\lambda_\alpha(x, y)$. It follows that if the field $\mathfrak{F}$ is taken sufficiently small, the values

$x, y, p(x, y), \lambda_\alpha(x, y)$ belonging to it will lie in so small a neighborhood of the sets (x, y, y', λ) belonging to E_{12} that the condition $II_{\mathfrak{N}}'$ will imply the inequality

$$E(x, y, p(x,y), \lambda(x,y), y') > 0$$

for every admissible set $(x, y, y') = (x, y, p)$ in $\mathfrak{F}$. Theorem 9:3 and the fundamental sufficiency theorem 7:2 now justify the theorem that was to be proved.

Bliss [IX, 736 - 7] has shown that if an extremal arc E_{12} satisfies the condition III' and is an extremal of a field $\mathfrak{F}$ with slope functions and multipliers $p_i(x, y)$, $\lambda_\alpha(x, y)$, then the inequality

$$E[x, y, p(x,y), \lambda(x,y), y'] > 0$$

holds for every admissible set $(x, y, y') \neq (x, y, p)$ in a neighborhood $\mathfrak{P}$ of the sets (x, y, y') on E_{12}. Hence by arguments like those in the preceding paragraph the following theorem is justified:

THEOREM 10:2. SUFFICIENT CONDITIONS FOR A WEAK RELATIVE MINIMUM. *If an admissible arc E_{12} satisfies all the conditions of the preceding theorem except the condition $II_{\mathfrak{N}}'$, then there are neighborhoods $\mathfrak{P}$ of the sets (x, y, y') on E_{12} and N of the end-values (x_1, y_1, x_2, y_2) of E_{12} such that the inequality $g(C_{34}) > g(E_{12})$ is true for every admissible arc C_{34} whose elements (x, y, y') are all in $\mathfrak{P}$, whose ends are in N and satisfy the conditions $\psi_\rho = 0$, and which is not*

identical with E_{12}.

Suppose now that the functions ψ_ρ are continuous at every pair of distinct or coincident points in a neighborhood of those belonging to E_{12}. Bliss has shown that if the ends of E_{12} are the only pair of distinct or coincident points on E_{12}, satisfying the conditions $\psi_\rho = 0$ then for every neighborhood N of the ends of E_{12} in $(x_1y_1x_2y_2)$-space there is a neighborhood $\mathfrak{F}$ of E_{12} in xy-space such that every pair of points (x_1, y_1), (x_2, y_2) in $\mathfrak{F}$ satisfying the conditions $\psi_\rho = 0$ are also in N [XII, p. 267]. Hence by suitably restricting the neighborhood $\mathfrak{F}$ of E_{12} in Theorem 10:1 we have the following corollary:

COROLLARY 10:1. *Let* E_{12} *be an admissible arc satisfying the conditions described in Theorem 10:1. If further the ends of* E_{12} *are the only pair of distinct or coincident points on* E_{12} *satisfying the conditions* $\psi_\rho = 0$, *then there is a neighborhood* $\mathfrak{F}$ *of* E_{12} *in xy-space such that the inequality* $g(C_{34}) > g(E_{12})$ *holds for every admissible arc* C_{34} *in* $\mathfrak{F}$ *with ends satisfying the conditions* $\psi_\rho = 0$ *and not identical with* E_{12}.

A similar corollary can be stated for weak relative minima.

BIBLIOGRAPHY

I. Kneser, Lehrbuch der Variationsrechnung, Braunschweig (1900), pp. 227 - 261.

II. Egorov, Die hinreichenden Bedingungen des Extremums in der Theorie des Mayerschen Problems, Mathematische Annalen, vol. 62 (1906), pp. 371 - 380.

III. Bolza, Über den anormalen Fall beim Lagrangeschen und Mayerschen Problem mit gemischten Bedingungen und variablen Endpunkten, Mathematische Annalen, vol. 74 (1913), pp. 430 - 446.

IV. Goursat, A Course in Mathematical Analysis, translated by Hedrick and Dunkel, vol. 2, Part 2., Ginn and Co. (1917).

V. Bliss, The problem of Mayer with variable end-points, Transactions of the American Mathematical Society, vol. 19 (1918), pp. 305 - 314.

VI. Larew, Necessary Conditions in the problem of Mayer in the calculus of variations, Transactions of the American Mathematical Society, vol. 20 (1919), pp. 1 - 22.

VII. Larew, The Hilbert integral and Mayer fields for the problem of Mayer in the calculus of variations, Transactions of the American Mathematical Society, vol. 26 (1924), pp. 61 - 67.

VIII. Kneser, Lehrbuch der Variationsrechnung (2), Braunschweig (1925), pp. 240 - 304.

IX. Bliss, The problem of Lagrange in the calculus of variations, American Journal of Mathematics, vol. 52 (1930), pp. 673 - 742.

X. Morse and Myers, The problems of Lagrange and Mayer with variable end-points. Proceedings of the American Academy of Arts and Sciences, vol. 66 (1931), pp. 235 - 253.

XI. Morse, Sufficient conditions in the problem of Lagrange

Mathematics, vol. 53 (1931), pp. 517 - 546.

XII. Bliss, The problem of Bolza in the calculus of variations, Annals of Mathematics, vol. 33 (1932), pp. 261 - 274.

SUFFICIENT CONDITIONS FOR THE GENERAL PROBLEM OF MAYER WITH VARIABLE END-POINTS

BY

MAGNUS RUDOLPH HESTENES

TABLE OF CONTENTS

SUFFICIENT CONDITIONS FOR THE GENERAL PROBLEM OF MAYER WITH VARIABLE END-POINTS[1]

1. Introduction. The problem of the calculus of variations to be considered here is the general problem of Mayer with variable end-points as proposed by Bliss (V, 305)[2] and recently studied for a particular case in a joint paper by Bliss and Hestenes (XVI, 1-39). As was remarked in the latter paper the general problem of Mayer is equivalent to the problem of Bolza, but the sets of sufficient conditions which have been given by Morse and Bliss for the problem of Bolza are not applicable to the problem of Mayer without further modification. In view of this fact it is the purpose of the present paper to establish a set of sufficient conditions for the general problem of Mayer with variable end-points. The proofs here given are equally applicable to the problem of Bolza considered as a problem of Mayer.

The procedure used is similar to that used by Bliss for the problem of Bolza (XII). We first derive in Section 4 a necessary condition analogous to that deduced by Bliss for the problem of Bolza. In Section 5 we construct an auxiliary problem of Mayer of the type discussed by Bliss and Hestenes. Their results are then applied in Sections 6 and 8 to the general problem by methods closely related to those of Mayer (XIII) and

---

[1]It is expected that this paper will appear in the Transactions of the American Mathematical Society, vol. 35.

[2]The Roman numerals refer to the bibliographies at the ends of the present and preceding papers. The Arabic numerals refer to pages.

Hahn (XIV, 127-136).

2. Statement of the problem. In the following pages the notation and the terminology used by Bliss and Hestenes for a particular problem of Mayer will be used throughout (XVI, 3, 4, 8). In addition it will be understood that the indices μ, ν have the ranges

$$\mu, \nu = 1, \cdots, p < 2n + 1.$$

The general problem of Mayer is then that of minimizing a function $g[x_1, y(x_1), x_2, y(x_2)]$ in a class of arcs

$$(2:1) \qquad y_i = y_i(x) \qquad (x_1 \leqq x \leqq x_2)$$

which satisfy the differential equations and end-conditions

$$\varphi_\alpha(x, y, y') = 0, \qquad \psi_\mu[x_1, y(x_1), x_2, y(x_2)] = 0.$$

As before the arcs (2:1) and the functions φ_α, g, ψ_μ will be assumed to have the continuity properties (a), (b), (c) (XVI, 4) in a neighborhood of a particular arc E_0 whose minimizing properties are to be studied, the determinant (2:1) appearing in (c) being now interpreted as a $(2n + 2)x(p + 1)$ dimensional matrix of rank $p + 1$.

For the general problem of Mayer the first necessary condition as given by Bliss and Hestenes (XVI, 5) is modified as follows, and is readily established by the methods which they suggest. The theorem has also been established by Morse and Myers (X, 245).

I. THE FIRST NECESSARY CONDITION. *Every minimizing arc E_0 for the problem of Mayer with variable end-points must satisfy, besides the conditions* (XVI, 5)

$$(2{:}2) \qquad F_{y_i'} = \int_{x_1}^{x} F_{y_i}\,dx + c_i, \qquad \varphi_\alpha = 0,$$

the further relation

$$(2{:}3) \qquad (F - y_i' F_{y_i'})dx + F_{y_i'}dy_i \Big|_1^2 + \lambda_0\, dg = 0$$

for every set of differentials dx_1, dy_{i1}, dx_2, dy_{i2} *satisfying the equations* $d\psi_\mu = 0$, λ_0 *being a suitably chosen constant.*

An admissible arc E_0 is said to be *normal relative to the end-conditions* $\psi_\mu = 0$ if there exist for it p sets of admissible variations ξ_1^ν, ξ_2^ν, $\eta_i^\nu(x)$ such that the determinant $\left|\Psi_\mu(\xi^\nu, \eta^\nu)\right|$ is different from zero (XVI, 6). For convenience an arc that is normal relative to the end-conditions $\psi_\mu = 0$ will be designated simply as *normal*.

THEOREM 2:1. *An admissible arc that does not satisfy the necessary condition I is normal.*

This follows at once because an admissible arc E_0 satisfies the necessary condition I if and only if every determinant of the form

$$\begin{vmatrix} G(\xi^\sigma, \eta^\sigma) \\ \Psi_\mu(\xi^\sigma, \eta^\sigma) \end{vmatrix} \qquad (\sigma = 1, \cdots, p + 1)$$

vanishes, where ξ_1^σ, ξ_2^σ, $\eta_i^\sigma(x)$ are p + 1 sets of admissible variations for E_0, and the function G is obtained from g in the same manner as Ψ_μ is obtained from ψ_μ (V, 309).

THEOREM 2:2. *An admissible arc* E_0 *that satisfies the necessary condition I is normal if and only if there exist for it no set of multipliers* $\lambda_\alpha(x)$, *not vanishing simultaneously, with which it satisfies equations* (2:2) *and for which all* (p + 1)-*rowed determinants of the matrix*

$$(2:4)\quad \left\|\begin{matrix} -y_{i1}'F_{y_i'}(x_1) & F_{y_i'}(x_1) & y_{i2}'F_{y_i'}(x_2) & -F_{y_i'}(x_2) \\ \Psi_{\mu x_1} & \Psi_{\mu y_{i1}} & \Psi_{\mu x_2} & \Psi_{\mu y_{i2}} \end{matrix}\right\|$$

vanish. If E_0 is normal the constant λ_0 can be chosen to be unity, the multipliers $\lambda_\alpha(x)$ with which E_0 satisfies the conditions (2:2) and (2:3) being then unique.

This theorem is an obvious generalization of a theorem given by Bliss and Hestenes and can be proved by the same methods (XVI, 6-7). A similar theorem has been established by Bolza (III, 441).

3. Theorems on extremals. It is known that in the problems of Mayer a non-singular extremal arc can be imbedded in a $(2n-1)$-parameter family of extremals (XVI, 12)

$$(3:1)\quad y_i = y_i(x, c_1, \cdots, c_{2n-1}), \quad \lambda_\alpha = \lambda_\alpha(x, c_1, \cdots, c_{2n-1}) \qquad (x_1 \leqq x \leqq x_2).$$

Further properties of this family are given in the following theorem:

THEOREM 3:1. Let E_0 be a member of the $(2n-1)$-parameter family of extremals (3:1) for parameter values (x_{10}, x_{20}, c_0). If the matrix

$$(3:2)\quad \left\|\begin{matrix} y_{ic_s}(x_1, c) \\ y_{ic_s}(x_2, c) \end{matrix}\right\|$$

has rank $2n-1$ on E_0, then there is a neighborhood N of the ends of E_0 in $(x_1y_1x_2y_2)$-space such that the end-values of every extremal of the family (3:1) with ends in N satisfy a relation $W(x_1, y_1, x_2, y_2) = 0$. Conversely, every pair of points (x_1, y_1), (x_2, y_2) in N satisfying the condition $W = 0$ can be joined by an

extremal E *of the family* (3:1), *and by taking* N *sufficiently small the parameters* (x_1, x_2, c) *belonging to* E *will lie in a preassigned* ϵ *-neighborhood of those belonging to* E_0. *The function* W *has continuous partial derivatives of the first two orders in* N.

The theorem can be proved as follows. Select 2n constants a_i, b_i such that the determinant

$$(3:3) \qquad \begin{vmatrix} y_{ic_s}(x_1, c) & a_i \\ y_{ic_s}(x_2, c) & b_i \end{vmatrix}$$

is different from zero on E_0. Consider now the equations

$$(3:4) \qquad \begin{aligned} y_{i1} &= y_i(x_1, c) + Wa_i, \\ y_{i2} &= y_i(x_2, c) + Wb_i. \end{aligned}$$

These equations are satisfied by the set $(x_{10}, y_{10}, x_{20}, y_{20}, c_0, W = 0)$ belonging to E_0. Furthermore the functional determinant with respect to the variables c_s, W is the determinant (3:3) and is therefore different from zero on E_0. It follows that equations (3:4) have a unique solution

$$(3:5) \qquad c_s = c_s(x_1, y_1, x_2, y_2), \qquad W = W(x_1, y_1, x_2, y_2)$$

in a neighborhood N of the end-values $(x_{10}, y_{10}, x_{20}, y_{20})$ belonging to E_0. The right members of equations (3:5) have continuous first and second derivatives in N since the right and left members of equations (3:4) have such derivatives. If now the end-values of an extremal are in N, then these end-values must satisfy the relation $W(x_1, y_1, x_2, y_2) = 0$ since the solutions of equations (3:4) are unique. Furthermore every set of values (x_1, y_1, x_2, y_2) in N satisfying the relation $W = 0$ are the end-

values of an extremal E with parameter values $[x_1, x_2, c(x_1, y_1, x_2, y_2)]$, and by taking N sufficiently small these parameter values will lie in a preassigned ϵ -neighborhood of those belonging to E_0. Hence the theorem is proved.

It is now possible to give an interesting geometric interpretation of normality.

THEOREM 3:2. <u>A non-singular extremal arc</u> E_0, <u>whose matrix</u> (3:2) <u>has rank</u> $2n - 1$, <u>is normal if and only if in the space of points</u> (x_1, y_1, x_2, y_2) <u>the extremal manifold</u> $W = 0$ <u>and the terminal manifold</u> $\psi_\mu = 0$ <u>are not tangent to each other at the point</u> $(x_{10}, y_{10}, x_{20}, y_{20})$ <u>defining the end-values of</u> E_0.

To prove this it is sufficient, as is readily seen, to show that the derivatives W_{x_1}, $W_{y_{i1}}$, W_{x_2}, $W_{y_{i2}}$ are proportional to the elements of the first row of the matrix (2:4). These derivatives have this property because the relation $F_{y_i'}\eta_i$ = constant along extremals (XVI, 5) with $\eta_i = y_{ic_s}dc_s$ implies that the differentials dx_1, dy_{i1}, dx_2, dy_{i2}, dc_s, dW belonging to equations (3:4) satisfy the relation

$$F_{y_i'}(dy_i - y_i'dx)\Big|_1^2 = F_{y_i'}y_{ic_s}dc_s\Big|_1^2 + h\,dW = h\,dW$$

where $h = b_1F_{y_1'}(x_2) - a_1F_{y_1'}(x_1)$. If $h = 0$ on E_0 then on account of the relation $F_{y_i'}\eta_i$ = constant, the determinant (3:3) would vanish on E_0 which is not the case. Hence $h \neq 0$ on E_0 and

$$(3:6)\qquad \begin{aligned} y_{i1}'F_{y_i'}(x_1) &= hW_{x_1}, & -F_{y_i'}(x_1) &= hW_{y_{i1}}, \\ -y_{i2}'F_{y_i'}(x_2) &= hW_{x_2}, & F_{y_i'}(x_2) &= hW_{y_{i2}} \end{aligned}$$

as was to be proved.

4. The necessary condition of Mayer. The necessary condition of Mayer for the problem of Bolza, as stated by Bliss (XII, 266), is also valid for the problem of Mayer considered here.[1] In order to derive this condition we suppose that E_0 is a normal non-singular minimizing arc without corners and hence, an extremal arc. If $\xi_1, \xi_2, \eta_i(x)$ is a set of admissible variations for E_0 which satisfy the conditions $\Psi_\mu(\xi, \eta) = 0$, then E_0 is a member of a one-parameter family of admissible arcs with ends satisfying the conditions $\psi_\mu = 0$ and having $\xi_1, \xi_2, \eta_i(x)$ as its variations along E_0 (IX, 695). For such a family the second variation of the function g to be minimized is expressible along E_0 in the form

$$(4:1) \qquad I_2 = (F_x + y_i' F_{y_i'})\xi^2 + 2F_{y_i'}\eta_i \xi \Big|_1^2 + 2(Q + \chi_\mu Q_\mu) + \int_{x_1}^{x_2} 2\omega(x, \eta, \eta')dx$$

where Q, Q_μ are quadratic forms in $\xi_1, \eta_i(x_1), \xi_2, \eta_i(x_2)$ whose coefficients are the second derivatives of the functions g, ψ_μ, respectively, and

$$2\omega(x, \eta, \eta') = F_{y_i y_k}\eta_i \eta_k + 2F_{y_i y_k'}\eta_i \eta_k' + F_{y_i' y_k'}\eta_i' \eta_k'.$$

This form for I_2 is readily obtained by the methods used by Bliss and Hestenes (XVI, 12-14). Let us consider variations satisfying the equations $\Psi_\mu(\xi, \eta) = 0$ along E_0, and of the special form $\xi_1 = dx_1$, $\xi_2 = dx_2$, $\eta_i = \delta y_i = y_{ic_s} dc_s$, where the functions $y_i(x, c)$ are those defining the $(2n - 1)$-parameter family (3:1) of extremals to which E_0 belongs. For such variations the second

[1] The proof is somewhat different from that given by Bliss for the problem of Bolza. He has called my attention to the fact that the argument which he used is inadequate in the case when the ends of E are conjugate, and has suggested the modifications indicated here.

variation (4:1) can also be expressed in the form

$$(4:2)\quad \begin{aligned} d^2g = {} & (F_x + y_i{}'F_{y_i})dx^2 + 2F_{y_i}\,\delta y_i\, dx + \\ & + \delta y_i\,\Omega_{\eta_i'}(x, \delta y, \delta y', \delta\lambda)\Big|_1^2 + 2(Q + \chi_\mu Q_\mu) \end{aligned}$$

given by Bliss (XII, 266), where $\delta\lambda_\alpha = \lambda_{\alpha c_s} dc_s$ and

$$\Omega(x, \eta, \eta', \mu) = \omega(x, \eta, \eta') + \mu_\alpha(\phi_{\alpha y_i}\eta_i + \phi_{\alpha y_i'}\eta_i').$$

Since E_0 is a minimizing arc the expression (4:1) must be ≥ 0 for all sets of admissible variations $\xi_1, \xi_2, \eta_i(x)$ which satisfy the conditions $\Psi_\mu(\xi, \eta) = 0$. In particular it must be ≥ 0 for variations $\xi_1 = dx_1$, $\xi_2 = dx_2$, $\eta_i = \delta y_i$ of the special type considered above satisfying the conditions $d\psi_\mu \equiv \Psi_\mu(dx, \delta y) = 0$. We have therefore the following result:

IV. THE NECESSARY CONDITION OF MAYER. *For a normal non-singular minimizing arc* E_0 *without corners the quadratic form (4:2) must satisfy the condition* $d^2g \geq 0$ *for all sets* $(dx_1, dx_2, dc_s) \neq (0,0,0)$ *which satisfy the equations* $d\psi_\mu = 0$. *Furthermore between the end-points* 1 *and* 2 *on* E_0 *there can be no point* 3 *conjugate to* 1 *defined by a value* x_3 *such that* E_0 *is normal on the interval* x_3x_2.

The last statement is the condition IV for problems of Mayer having 2n + 1 end-conditions (XVI, 19), valid here for E_0 since E_0 must also be a minimizing arc for such a problem, as will be seen in the next section.

5. *An auxiliary problem of Mayer.* In order to construct a problem of Mayer of the type described in the last paragraph we suppose that E_0 is a minimizing arc for the general problem of Mayer considered here. Its end values $(x_{10}, y_{10}, x_{20}, y_{20})$

satisfy the conditions $\psi_\mu = 0$ ($\mu = 1, \cdots, p$). Adjoin to the functions ψ_μ, $2n + 1 - p$ functions $\psi_\tau(x_1, y_1, x_2, y_2)$ ($\tau = p + 1, \cdots, 2n + 1$) possessing continuous first and second partial derivatives in a neighborhood of the values ($x_{10}, y_{10}, x_{20}, y_{20}$), vanishing at these values, and having the determinant

$$(5:1) \qquad \begin{vmatrix} g_{x_1} & g_{y_{i1}} & g_{x_2} & g_{y_{i2}} \\ \psi_{\rho x_1} & \psi_{\rho y_{i1}} & \psi_{\rho x_2} & \psi_{\rho y_{i2}} \end{vmatrix}$$

different from zero on E_0. The new set of end-conditions $\psi_\rho = 0$ ($\rho = 1, \cdots, 2n + 1$) defines an auxiliary problem of Mayer of the type discussed by Bliss and Hestenes. It is clear that E_0 is also a minimizing arc for this auxiliary problem.

THEOREM 5:1. <u>Let E_0 be an admissible arc that is normal on the interval $x_{10}x_{20}$ and satisfies the necessary condition I. If E_0 is normal relative to the end-conditions $\psi_\mu = 0$ ($\mu = 1, \cdots, p$), then it is normal relative to the end-conditions $\psi_\rho = 0$ ($\rho = 1, \cdots, 2n + 1$) just defined.</u>

To prove this theorem we recall that the matrix (2:4) has rank $p + 1$ since E_0 is normal relative to the end-conditions $\psi_\mu = 0$. Furthermore since E_0 satisfies the transversality condition (2:3), it follows that on E_0 the derivatives g_{x_1}, $g_{y_{i1}}$, g_{x_2}, $g_{y_{i2}}$ are expressible as a linear combination of the rows of the matrix (2:4), the multiplier of the first row being different from zero. The rank of the matrix (2:4) formed for the new end-conditions $\psi_\rho = 0$ is therefore unaltered when the elements of the first row are replaced by the derivatives g_{x_1}, $g_{y_{i1}}$, g_{x_2}, $g_{y_{i2}}$. The matrix thus formed is the matrix of the determinant (5:1) and has rank $2n + 2$. Hence according to Theorem 2:2, E_0 is also normal relative to the end-conditions $\psi_\rho = 0$, and the

theorem is established.

6. *A fundamental sufficiency theorem.* With the help of the auxiliary problem just constructed we can prove the following theorem:

THEOREM 6:1. A FUNDAMENTAL SUFFICIENCY THEOREM. *Let* E_0 *be an extremal arc with the following properties:*

(A) E_0 *satisfies the sufficient conditions for a proper strong relative minimum with respect to admissible arcs* C *satisfying the end-conditions* $\psi_\rho(C) = 0$ *of the auxiliary problem of Mayer defined in Section* 5.

(B) *There is a neighborhood* M *of the ends of* E_0 *in* $(x_1y_1x_2y_2)$*-space such that the inequality* $g(E) > g(E_0)$ *holds for every extremal* E *of the family* (3:1) *with ends in* M *satisfying the conditions* $\psi_\mu(E) = 0$ *and not identical with* E_0.

Then there exist neighborhoods $\mathfrak{F}$ *of* E_0 *in* xy*-space and* N *of the ends of* E_0 *in* $(x_1y_1x_2y_2)$*-space such that the inequality* $g(C) > g(E_0)$ *holds for every admissible arc* C *in* $\mathfrak{F}$ *with ends in* N *satisfying the conditions* $\psi_\mu(C) = 0$ *and not identical with* E_0.

The proof is based on the following two lemmas, the proofs of which will be given in the next section.

LEMMA 6:1. (*Modification of Hahn's Theorem* (XIV, 129)). *The property (A) for* E_0 *implies the existence of neighborhoods* $\mathfrak{F}$ *of* E_0 *in* xy*-space and* M *of the ends of* E_0 *in* $(x_1y_1x_2y_2)$*-space such that for every extremal* E *of the family* (3:1) *with ends in* M *the inequality* $g(C) > g(E)$ *holds for every admissible arc* C *in* $\mathfrak{F}$ *with ends in* M *satisfying the conditions* $\psi_\rho(C) = \psi_\rho(E)$ *and not identical with* E.

LEMMA 6:2. *The property (A) for* E_0 *implies that every neighborhood* M *of the end-values of* E_0 *has associated with it a*

<u>second neighborhood</u> N <u>of these end-values such that for every admissible arc</u> C <u>with ends in</u> N <u>there is an extremal</u> E <u>of the family</u> (3:1) <u>with ends in</u> M <u>satisfying the conditions</u> $\psi_\rho(C) = \psi_\rho(E)$.

With the help of these lemmas the proof of Theorem 6:1 is as follows. Select first neighborhoods $\mathcal{F}$ of E_0 and M of the ends of E_0 effective as in Lemma 6:1 and as in (B). Select a second neighborhood N of the ends of E_0 related to M as in Lemma 6:2. Consider now an admissible arc C in $\mathcal{F}$ with ends in N satisfying the conditions $\psi_\mu = 0$. According to Lemma 6:2 there is an extremal E of the family (3:1) with ends in M satisfying the conditions $\psi_\mu(E) = 0$, $\psi_\tau(E) = \psi_\tau(C)$, where the functions ψ_τ are those adjoined to the functions ψ_μ to form the auxiliary Mayer problem defined in Section 5. From Lemma 6:1 it follows that $g(C) \geqq g(E)$, and from the property (B) we have $g(E) \geqq g(E_0)$. Hence $g(C) \geqq g(E_0)$, the equality being valid only in case C coincides with E_0 as was to be proved.

7. <u>Proofs of two lemmas</u>. In order to prove Lemma 6:1 we use the result obtained by Bliss and Hestenes (XVI, 32-33) which states that the property (A) for E_0 given in Theorem 6:1 implies the existence of a function $W(a_1 \cdots a_n)$ such that the n-parameter family of extremals

$$(7{:}1)\quad y_i = y_i(x, x_{20}, a, W_a), \quad z_i = z_i(x, x_{20}, a, W_a) \quad (x_1 \leqq x \leqq x_2)$$

contains E_0 for parameter values (x_{10}, x_{20}, a_0) and has the determinant $|y_{ia_k}|$ different from zero along E_0. Furthermore each extremal E of the family (7:1) has on it the element $(x, y, z) = (x_{20}, a, W_a)$, where the a_i are the parameter values defining E. If now we select n - 1 functions $W_r(a_1, \cdots, a_n)$ having continuous

first and second partial derivatives. and such that the determinant $|W_{a_i} \quad W_{ra_i}|$ is different from zero for the values $a_i = a_{i0}$, then the $(2n - 1)$-parameter family of extremals

$$
\begin{aligned}
y_i &= y_i(x, x_{20}, a, W_a + b_r W_{ra}) = y_i(x, a, b), \\
z_i &= z_i(x, x_{20}, a, W_a + b_r W_{ra}) = z_i(x, a, b)
\end{aligned} \tag{7:2}
$$

$$(x_1 \leqq x \leqq x_2)$$

contains E_0 for parameter values $(x_{10}, x_{20}, a_0, b = 0)$. Moreover every extremal E of this family has on it the element $(x, y_i, z_i) = (x_{20}, a_i, W_{a_i} + b_r W_{ra_i})$, where the parameter values a_r, b_r are those defining E. The equations expressing this fact are the equations

$$a_i = y_i(x_{20}, a, b), \qquad W_{a_i} + b_r W_{ra_i} = z_i(x_{20}, a, b),$$

and by differentiation it is found that the determinant

$$\begin{vmatrix} y_{ia_k} & y_{ib_r} & 0 \\ z_{ia_k} & z_{ib_r} & z_i \end{vmatrix}$$

is different from zero for the values $(x, a, b) = (x_{20}, a_0, 0)$. Hence the family (7:2) is one of the type (3:1), its multipliers $\lambda_\alpha(x, a, b)$ being found in the usual manner (XVI, 8-12).

Since the determinant $|\, y_{ia_k} \,|$ belonging to the family (7:1) is different from zero on E_0, the determinant $|y_{ia_k}(x,a,b)|$ belonging to the family (7:2) has the same property. Hence the system of equations

$$y_i = y_i(x, a, b) \tag{7:3}$$

has a unique solution

$$a_i = a_i(x, y, b)$$

in a neighborhood $\mathfrak{D}$ of the values (x, y, b) belonging to E_0. The functions $a_i(x, y, b)$ are continuous and possess continuous derivatives of the first two orders in the domain $\mathfrak{D}$. If now we let

$$(7:4)\qquad \begin{aligned} p_i(x, y, b) &= y_{ix}[x, a(x, y, b), b] , \\ \lambda_\alpha(x, y, b) &= \lambda_\alpha[x, a(x, y, b), b] , \end{aligned}$$

then according to the condition II'_n implied by the property (A) on E_0, the domain $\mathfrak{D}$ can be so restricted that at each element (x, y, b) in $\mathfrak{D}$ the inequality

$$E[x, y, p(x, y, b), \lambda (x, y, b), y'] > 0$$

holds for every admissible set $(x, y, y') \neq (x, y, p)$, where $E(x, y, p, \lambda , y')$ is the Weierstrass E-function (XVI, 24, 35). Furthermore on the hyperplane $x = x_{20}$ in xy-space the Hilbert integral I^* is independent of the path (XVI, 33; cf. XII, 269) when the parameters b_r are fixed. It follows that for each set b_r the region $\mathfrak{F}$ of points (x, y), whose elements (x, y, b) are all in $\mathfrak{D}$, forms a field with slope functions and multipliers defined by equations (7:4) (XVI, 31-32). We have a family of such fields depending upon the $n - 1$ parameters b_r. In each field the Weierstrass E-function is > 0 unless $y_i' = p_i$. Hence according to a theorem proved by Bliss and Hestenes (XVI, 26) there is a neighborhood M of the end-values of E_0 such that every extremal E with ends in M and belonging to one of these fields furnishes a proper strong relative minimum for the function g in the class of admissible arcs C in $\mathfrak{F}$ whose ends are in M and satisfy the conditions $\psi_\rho (C) = \psi_\rho (E)$.

Lemma 6:1 will now be established completely if we show that the neighborhood M of the ends of E_0 can be restricted so that every extremal E of the family (7:2) with ends in M is a member of one of the fields just described. To do this we select a constant h so that the set $[x, y, a_1(x, y, b), b]$ with elements (x, y, b) in $\mathfrak{D}$ is the only solution of equations (7:3) satisfying the relation

$$(7:5) \qquad a_1(x, y, b) - h \leqq a_1 \leqq a_1(x, y, b) + h.$$

This can always be done since the solution $a_1(x, y, b)$ of equations (7:3) is isolated. We now select a constant ϵ such that the inequalities

$$|a_1 - a_{1o}| < h/2, \qquad |a_{1o} - a_1(x, y, b)| < h/2$$

hold along every extremal E of the family (7:2) with parameter values (x_1, x_2, a, b) in an ϵ-neighborhood of those belonging to E_0. The relation (7:5) now holds for every set of values (x, y, a, b) on E. It follows that $a_1 = a_1(x, y, b)$, and hence E is an extremal of one of the fields just described. This completes the proof of Lemma 6:1 since according to Theorem 3:1 the neighborhood M of the ends of E_0 can be so restricted that every extremal E of the family (7:2) with ends in M has parameter values (x_1, x_2, a, b) in the ϵ-neighborhood just defined.

In order to prove Lemma 6:2 consider first the equations

$$(7:6) \qquad W(x_1, y_1, x_2, y_2) = 0, \qquad \psi_\rho(x_1, y_1, x_2, y_2) = m_\rho ,$$

where W is the function defined in Theorem 3:1. As was seen in Section 3, the functional determinant of these equations is different from zero on E_0. Furthermore equations (7:6) are satis-

fied by the set $(x_1, y_1, x_2, y_2, m) = (x_{10}, y_{10}, x_{20}, y_{20}, 0)$ belonging to E_0. Hence there is a constant $h > 0$ such that equations (7:6) have a unique solution

$$(7{:}7) \qquad x_1 = x_1(m), \quad y_{11} = y_{11}(m), \quad x_2 = x_2(m), \quad y_{12} = y_{12}(m)$$

for all values m_ρ satisfying the relations $|m_\rho| < h$. If h is sufficiently small, then according to Theorem 3:1 every pair of points (x_1, y_1), (x_2, y_2) can be joined by an extremal of the family (3:1). Furthermore it is clear that, if necessary, the constant h can be further restricted so that every set of values (x_1, y_1, x_2, y_2) defined by equations (7:7) with $|m_\rho| < h$ is in a preassigned neighborhood M of the end-values of E_0. If now we select a second neighborhood N of the end-values of E_0 so that every set of values (x_1, y_1, x_2, y_2) in N satisfies the relation $|\psi_\rho(x_1, y_1, x_2, y_2)| < h$, then every admissible arc C with ends in N determines a set of values $m_\rho = \psi_\rho(C)$ satisfying the relation $|m_\rho| < h$, and these in turn determine an extremal arc E with ends in M satisfying the conditions $\psi_\rho(E) = \psi_\rho(C)$. This proves Lemma 6:2.

8. Sufficient conditions for relative minima. The necessary condition I is given in Section 2. The symbols II_N', III' will be used to denote the strengthened conditions of Weierstrass and Clebsch as defined by Bliss and Hestenes (XVI, 34-35). The symbol IV' will be used to denote the condition IV of Section 4 strengthened so as to exclude the equality sign. With these definitions agreed upon we can state the following theorem:

THEOREM 8:1. SUFFICIENT CONDITIONS FOR A STRONG RELATIVE MINIMUM. *Let E_0 be an admissible arc without corners and with end-points determined by values x_{10}, x_{20} and satisfying the*

conditions $\psi_\mu = 0$. *If* E_0 *is normal relative to the end-conditions* $\psi_\mu = 0$, *is normal on every sub-interval* $x_{10}x_3$ *of* $x_{10}x_{20}$, *and satisfies the conditions* I, II'_n, III′, IV′, *then there exist neighborhoods* $\mathfrak{F}$ *of* E_0 *in xy-space and* N *of the ends of* E_0 *in* $(x_1y_1x_2y_2)$*-space such that the inequality* $g(C) > g(E_0)$ *holds for every admissible arc* C *in* $\mathfrak{F}$ *with ends in* N *satisfying the conditions* $\psi_\mu(C) = 0$ *and not identical with* E_0.

The theorem will be established if we can show that the hypotheses of the theorem imply those of Theorem 6:1. It is easily seen from Theorem 5:1 and from the sufficiency conditions given by Bliss and Hestenes for the case $p = 2n + 1$ (XVI, 35) that E_0 is an extremal arc having the property (A) of Theorem 6:1 provided that we can show that the condition IV′ implies that the ends of E_0 are not conjugate to each other. If the ends of E_0 were conjugate then the constants dc_s in the expressions $\delta y_i = y_{ic_s} dc_s$ could be selected, not all zero, so that the differentials δy_i would all vanish at the ends of E_0. If we should take these constants dc_s together with the values $dx_1 = dx_2 = 0$, then the condi- $d\psi_\mu = 0$ would be satisfied and the expression (4:1) for d^2g would vanish, which would contradict the condition IV′. Hence E_0 has property (A) of Theorem 6:1.

To prove that E_0 has the property (B) of Theorem 6:1, we first note that the conditions I, III′ imply the existence of a family of extremals (3:1) containing E_0 for parameter values (x_{10}, x_{20}, c_{s0}). From conditions I, IV′ it follows that $dg = 0$, $d^2g > 0$ for every set of differentials $(dx_1, dx_2, dc_s) \neq (0,0,0)$ which satisfy the conditions $d\psi_\mu = 0$. But these are the conditions which insure that $g(x_1, x_2, c_s) > g(x_{10}, x_{20}, c_{s0})$ for all sets $(x_1, x_2, c_s) \neq (x_{1}, x_2, c_{s0})$ satisfying the equations $\psi_\mu(x_1, x_2, c_s) = 0$ and lying in a sufficiently small ϵ-neigh-

borhood of (x_{10}, x_{20}, c_{s0}). Furthermore since the ends of E_0 are not conjugate the matrix (3·2) has rank 2n - 1 (XVI, 21), and according to Theorem 3:1 there is a neighborhood M of the ends of E_0 such that every extremal with ends in M has parameter values (x_1, x_2, c_s) in the ϵ -neighborhood described above. It follows that $g(x_1, y_1, x_2, y_2) > g(x_{10}, y_{10}, x_{20}, y_{20})$ for every extremal E with ends in M satisfying the conditions $\psi_\mu(E) = 0$ and not identical with E_0. Hence E_0 has the property (B) and the theorem is established.

In a similar manner sufficient conditions for a weak relative minimum for the general problem of Mayer with variable end-points can be established. The argument is like that of Bliss and Hestenes(XVI, 36) with the help of simple modifications of Theorem 6:1 and Lemma 6:1 above. The Theorem 10:2 of Bliss and Hestenes remains valid here if we replace the phrase "preceding theorem" by "Theorem 8:1" and the equations $\psi_\rho = 0$ by $\psi_\mu = 0$. Similarly Corollary 10:1 of the paper by Bliss and Hestenes is still effective if we replace "Theorem 10:1" by "Theorem 8:1" and ψ_ρ by ψ_μ.

BIBLIOGRAPHY

The papers listed below are a continuation of the list at the end of the paper of Bliss and Hestenes cited here as XVI.

XIII. Mayer, Zur Aufstellung der Kriterion des Maximums und Minimums der einfachen Integrale bei variablen Grenzwerten, Leipziger Berichte, vol. 36 (1884), pp. 99-128, vol. 48 (1896), pp. 436-465.

XIV. Hahn, Ueber Variationsprobleme mit variablen Endpunkten, Monatshefte für Mathematik und Physik, vol. 22 (1911), pp. 127-136.

XV. Hancock, Theory of Maxima and Minima, Ginn and Company, 1917.

XVI. Bliss and Hestenes, Sufficient Condition for a Problem of Mayer in the Calculus of Variations,

THE PROBLEM OF BOLZA AND ITS ACCESSORY BOUNDARY VALUE PROBLEM

BY

KUEN-SEN HU

TABLE OF CONTENTS.

THE PROBLEM OF BOLZA AND ITS
ACCESSORY BOUNDARY VALUE PROBLEM

INTRODUCTION

The problem of Lagrange with variable end points and in a generalized form has been formulated by Bolza.[1] He considered the problem of minimizing an expression of the form

$$g[x_1, y_i(x_1), x_2, y_i(x_2)] + \int_{x_1}^{x_2} f(x, y_i, y_i')\,dx, \qquad (i = 1, \cdots, n)$$

with respect to certain arcs which satisfy a set of finite and differential equations and also certain general end conditions. Bliss[2] lightened Bolza's hypotheses and put the problem in a form in which the integral does not appear. In this way his problem is most symmetrical and forms a direct generalization of the classical problem of Mayer. Morse[3] recently considered the problem of Bolza and gave for the first time a set of sufficient conditions for it. The conditions make use of the characteristic numbers of a boundary value problem which arises from minimizing the second variation for the general problem. He attacked this problem by means of the theory of what he called a fundamental quadratic form.

---

1 Über den "Anormalen Fall" beim Lagrangeschen und Mayerschen Problem mit gemischten Bedingungen und variablen Endpunkten. Mathematische Analen. Vol. 74 (1913). p. 430.

2 The problem of Mayer with variable end points. Transactions of the American Mathematical Society. Vol. 19 (1918), pp. 305 - 314.

3 Sufficient conditions in the problem of Lagrange with variable end points. American Journal of Mathematics. Vol. 53 (1931). pp. 517 - 546.

In this dissertation, as suggested by Professor Bliss, we have discussed this boundary value problem purely from the theory of differential equations.[1] By a simple change of notation we are able to avoid the non-tangency condition imposed by Morse and at the same time put the problem in its most symmetrical form. By first proving various properties of the characteristic numbers and functions we have arrived at an important expansion theorem which enables us to show that a necessary and sufficient condition for the second variation to be always positive (or non-negative) is that the characteristic numbers should be all positive (or non-negative). We have also attached a minimizing property to each of the characteristic numbers and functions which corresponds to the complete problem of Morse. For the simplest plane problem a similar proof has been given by Hilbert[2] as an application of the theory of linear integral equations with symmetric kernels. In our case the expansion theorem is also deducible[3] from the theory of a system of linear integral equations with symmetric kernels but the method which we have adopted is more direct and elementary.

By the ordinary differentiation method, a generalization of Hahn's lemma, and the condition for the second variation to

---

[1] In this connection the procedures of Bliss in the paper "A boundary value problem for a system of ordinary linear differential equations of the first order" Transactions of the American Mathematical Society. Vol. 28 (1926). pp. 561 - 584 have been of much assistance.

[2] Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen, 1912, p. 56. Also Lovitt, "Linear integral equations," 1924, p. 196.

[3] See Lemma 8·1 (below) and the remark following its proof.

be always positive, a new sufficiency theorem has been given for the problem of Bolza independent of Morse and without the non-tangency condition. Instead of assuming the minimizing arc E_{12} to be normal on every subinterval of x_1x_2 relative to the Euler-Lagrange equations, we assume that it is normal only on every subinterval x_1x_3. Both of these normality conditions require the integrand function f to be not identically zero. In this respect the problem of Bolza may be regarded as completely solved but for the problem of Bliss a somewhat different kind of normality condition must be needed.

Finally by considering various boundary value systems in connection with their corresponding minimizing problems we have deduced some rather general oscillation and comparison theorems which are involved with the theory of conjugate and focal points. Our results include those of Morse in the paper already cited and in one of his earlier papers.[1] The proofs are also quite independent and seem to be easier. A fundamental lemma[2] which leads to the construction of a field surrounding an extremal which is only normal on every subinterval or on every subinterval x_1x_3 also comes out as an application of a more general comparison theorem.

---

1. A generalization of the Sturm separation and comparison theorems in n-space, Mathematische Annalen. Vol. 103 (1930), pp. 52-69.

2 See Theorem 14·2 (below). A proof has been given by Bliss in class lectures, Spring quarter, 1931. Another by Morse, "Sufficient conditions in the problem of Lagrange with fixed end points," Annals of Mathematics. Vol. 32 (1931), pp. 574, 576.

CHAPTER I.

THE PROBLEMS AND SOME OF THEIR NECESSARY CONDITIONS.

1. The problem of Bolza and its multiplier rule. By definition let an admissible arc be an arc

$$y_i = y_i(x), \qquad (x_1 \leq x \leq x_2;\ i = 1, \cdots, n) \tag{1·1}$$

of class D' lying in an open region $\mathcal{R}$ of the (x,y,y')-space and satisfying a set of differential equations

$$\varphi_\alpha(x, y, y') = \varphi_\alpha(x, y_i, y_i') = 0, \qquad (\alpha = 1, \cdots, m < n) \tag{1·2}$$

where y_i' indicates the derivative of y_i with respect to x. The problem of Bolza is then the following

Problem 1. To find in the class of arcs $y_i = y_i(x)$, admissible according to the above definition and satisfying a set of end conditions

$$\psi_\mu(x, y) = \psi_\mu(x_1, y_{i1}, x_2, y_{i2}) = 0,^{1} \qquad (\mu = 1, \cdots, p \leq 2n+2) \tag{1·3}$$

one which minimizes an expression of the form

$$I = g(x, y) + \int_{x_1}^{x_2} f(x, y, y')\,dx = g(x_1, y_{i1}, x_2, y_{i2}) + \int_{x_1}^{x_2} f(x, y_i, y_i')\,dx. \tag{1·4}$$

As usual we concentrate attention on a particular admissible arc E_{12} with equations of the form (1·1) and satisfying the end conditions (1·3). We investigate what properties it must have in order to be a minimizing arc. The analysis is

---

[1] The notations y_{i1} and y_{i2} represent the values of y_i at x_1 and x_2 respectively. In the sequel similar notations will be used for other functions, like η_{i1}, ζ_{i1}, etc.

based upon the following hypotheses:

(a) The functions φ_α and f are of class C''' for values of their arguments in the region $\mathcal{R}$ while the functions ψ_μ and g need only be of class C'' for x_1, y_{i1}, x_2, y_{i2} in the projection of $\mathcal{R}$;

(b) The m by n matrix $\| \varphi_{\alpha y_i'} \|$ is of rank m at every point of E_{12};

(c) The p by 2n - 2 matrx $\| \psi_{\mu x_1}, \psi_{\mu y_{i1}}, \psi_{\mu x_2}, \psi_{\mu y_{i2}} \|$ has rank p for end values of E_{12}.

Here and in the following pages literal subscripts, following the indices of functions and elsewhere, like y_i', y_{i1}, y_{i2} in (b) and (c), will be frequently used to indicate partial derivatives.

We shall now give some definitions and the multiplier rule theorem without proof. Let

$$(1\cdot5) \qquad y_i = y_i(x, b), \qquad (x_1(b) \leq x \leq x_2(b))$$

be a one-parameter family of admissible arcs containing E_{12} for b = 0 and of class C' in b near b = 0. By differentiating the equations $\varphi_\alpha[x, y(x,b), y'(x,b)] = 0$ with respect to b and then putting b = 0, we shall have the so-called <u>equations of variation along E_{12} for the functions φ_α</u>

$$(1\cdot6) \qquad \Phi_\alpha(x,\eta,\eta') = \varphi_{\alpha y_i'}\eta_i' + \varphi_{\alpha y_i}\eta_i = 0 \text{ [1]},$$

where η_i denotes $y_{ib}(x,0)$ and the arguments in the partial derivatives of φ_α are those belonging to E_{12}. If the family (1·5) also satisfies the end conditions (1·3) for every b near

---

1 Throughout this paper a repetition of subscripts will indicate sum as in tensor analysis, e.g. $\varphi_{\alpha y_i}\eta_i = \varphi_{\alpha y_1}\eta_1 + \cdots + \varphi_{\alpha y_n}\eta_n$.

$b = 0$, then by differentiating the equations $\psi_\mu[x(b), y(b)] = 0$ with respect to b and setting $b = 0$, we shall have

$$(1\cdot 7) \qquad \Psi_\mu(\xi,\eta) = (\psi_{\mu x_1} + \psi_{\mu y_{i1}} y'_{i1})\xi_1 + \psi_{\mu y_{i1}}\eta_{i1} + (\psi_{\mu x_2} + \psi_{\mu y_{i2}} y'_{i2})\xi_2 + \psi_{\mu y_{i2}}\eta_{i2} = 0 .$$

where $\xi_1 = x_{1b}(0)$, $\xi_2 = x_{2b}(0)$ and the arguments in the partial derivatives of ψ_μ are those belonging to the ends of E_{12}. These are called the <u>equations of variation along</u> E_{12} <u>for the functions</u> ψ_μ.

A set of <u>admissible variations</u> along E_{12} is by definition a set ξ_1, ξ_2, $\eta_i(x)$ in which ξ_1, ξ_2 are arbitrary constants and $\eta_i(x)$ form a solution of class D' of the equations of variation (1·6). The arc E_{12} will be called <u>normal</u> if there exist for it p sets of admissible variations ξ_1^ν, ξ_2^ν, $\eta_i^\nu(x)$, $(\nu = 1, \cdots, p)$ such that

$$(1\cdot 8) \qquad \left| \Psi_\mu(\xi^\nu, \eta^\nu) \right| \neq 0 , \qquad (\mu, \nu = 1, \cdots, p).$$

This kind of normality is normality relative to the conditions $\psi_\mu = 0$. Later on in the sufficiency proof we shall need the following kind of stronger normality condition. The arc E_{12} is called <u>absolutely normal</u> on a subinterval x_3x_4 of x_1x_2 if there exist for it 2n sets of admissible variations $\eta_i^\tau(x)$, $(\tau = 1, \cdots, 2n)$ such that

$$(1\cdot 9) \qquad \left| \eta_i^\tau(x_3), \ \eta_i^\tau(x_4) \right| \neq 0 , \qquad (i = 1, \cdots, n; \ \tau = 1, \cdots, 2n).$$

The adverb "absolutely" is introduced here for the first time. Its use seems to be justified because, if E_{12} is absolutely normal on a certain interval, it will be normal on the same interval relative to any conditions.

Theorem 1·1.[1] *For every minimizing arc* E_{12} *of Problem* 1, *there exists a set of constants* c_i *and a function*

$$F(x,y,y',\lambda) = \lambda_0 f + \lambda_\alpha(x)\varphi_\alpha \tag{1·10}$$

such that the equations

$$F_{y_i'} = \int_{x_1}^{x} F_{y_i}\,dx + c_i \tag{1·11}$$

are satisfied at every point of E_{12}. *The constant* λ_0 *and the multipliers* $\lambda_\alpha(x)$ *do not all vanish at any point of* x_1x_2 *and are continuous except possibly at values of* x *defining corners of* E_{12}. *Furthermore there exist constants* d_μ *such that* λ_0 *and* d_μ *are not all zero and satisfy with the end values of* E_{12} *the so-called transversality conditions*

$$\begin{aligned} -F(x_1) + y_{i1}' F_{y_i'}(x_1) + \lambda_0 g_{x_1} + d_\mu \psi_{\mu x_1} &= 0, \\ F(x_2) - y_{i2}' F_{y_i'}(x_2) + \lambda_0 g_{x_2} + d_\mu \psi_{\mu x_2} &= 0, \\ -F_{y_i'}(x_1) + \lambda_0 g_{y_{i1}} + d_\mu \psi_{\mu y_{i1}} &= 0, \\ F_{y_i'}(x_2) + \lambda_0 g_{y_{i2}} + d_\mu \psi_{\mu y_{i2}} &= 0, \end{aligned} \tag{1·12}$$

where $F(x_1)$ *represents the value of* F *at* $x = x_1$ *and* $F(x_2)$, $F_{y_i'}(x_1)$, $F_{y_i'}(x_2)$ *have similar meanings. For a normal minimizing arc* E_{12}, *the constant* λ_0 *cannot be zero, and the multipliers and constants taken in the form* $\lambda_0 = 1$, $\lambda_\alpha(x)$ *and* c_i,

---

[1] This theorem is a consequence of the methods used by Bliss in "*The problem of Lagrange in the calculus of variations*" American Journal of Mathematics, Vol. 52 (1930), p. 692 and in "*The problem of Mayer*" loc. cit., p. 309 . An explicit proof has been given by Morse and Myers, *The problems of Lagrange and Mayer with variable end-points*, Proceedings of the American Academy of Arts and Sciences, Vol. 66 (1931), p. 245.

d_μ __are__ __unique.__

__Corollary__ 1. __At every point of__ E_{12}, __not a corner, the Euler-Lagrange equations__

$$(1\cdot13)\qquad (x,y,y')=0\,,\qquad (d/dx)F_{y_i'}=F_{y_i}$$

__must be satisfied.__

__Corollary__ 2. __At every corner of__ E_{12} __the corner conditions__

$$(1\cdot14)\qquad F_{y_i'}[x,y,y'(x-0),\lambda(x-0)]=F_{y_i'}[x,y,y'(x+0),\lambda(x+0)]$$

__must be satisfied.__

__Corollary__ 3. __Near every point of__ E_{12} __which is not a corner and at which__

$$(1\cdot15)\qquad R=\begin{vmatrix} F_{y_i'y_k'} & \varphi_{\alpha y_i'} \\ \varphi_{\alpha y_k'} & 0 \end{vmatrix} \qquad \left(\begin{matrix} i,k=1,\cdots,n \\ \alpha=1,\cdots,m \end{matrix}\right)$$

__is different from zero, the arc__ E_{12} __is of class__ C" __and the multipliers__ $\lambda_\alpha(x)$ __are of class__ C'.

An admissible arc $y_i = y_i(x)$ with a set of multipliers λ_o, $\lambda_\alpha(x)$ is called __an extremal__, if $y_i(x)$ are of class C" and $\lambda_\alpha(x)$ are of class C' and if further they satisfy the Euler-Lagrange equations (1·13). A minimizing arc or an extremal is called __non-singular__ if the determinant in (1·15) is different from zero at every point of it. If E_{12} is a non-singular minimizing arc for Problem 1 and has no corners, then by Corollary 3 it must be an extremal. In order to avoid repetition of statements in the following discussions we add to the hypotheses (a), (b), (c) already made the assumption:

(d) E_{12} is a normal non-singular extremal arc satis-

fying the end conditions $\psi_\mu = 0$ and the transversality conditions (1·12).

2. <u>The second variation for E_{12} and its accessory problem</u>. Since E_{12} is a normal extremal arc satisfying the conditions $\psi_\mu = 0$, it follows that, for every set of admissible variations ξ_1, ξ_2, $\eta_i(x)$ satisfying the conditions $\Psi_\mu(\xi,\eta) = 0$, there exists[1] a one-parameter family of admissible arcs of the form (1·5), satisfying the conditions $\psi_\mu = 0$, containing E_{12} for $b = 0$, of class C'' in b near $b = 0$, and having ξ_1, ξ_2, $\eta_i(x)$ as its variations along E_{12}. When this family is substituted in the expression for I in (1·4), we shall have a function $I(b)$. If we take $\lambda_0 = 1$ in theorem 1·1, the <u>first variation</u> along E_{12} for this family is readily verified to be

$$
\begin{aligned}
& I_1(\xi,\eta) = (d/db)\, I(b)\Big|^{b=0} \\
(2\cdot1) \qquad & \quad = G(\xi,\eta) + F(x_2)\xi_2 - F(x_1)\xi_1 + \int_{x_1}^{x_2} (F_{y_i}\eta_i + F_{y_i'}\eta_i')\,dx,
\end{aligned}
$$

where $G(\xi,\eta)$ is given by (1·7) with ψ_μ replaced by g, and x_1, x_2 refer to the ends of E_{12}. The value of $I_1(\xi,\eta)$ is easily seen to be zero by means of the equations (1·11) and the transversality conditions (1·12).

With the help of (1·11), (1·12) and some complicated but not difficult computations, we can also verify that the <u>second variation</u> along E_{12} for the above mentioned family has the form[2]

$$
(2\cdot2) \qquad I_2(\xi,\eta) = (d^2/db^2)\, I(b)\Big|^{b=0} = 2q(\xi,\eta) + \int_{x_1}^{x_2} 2\omega(x,\eta,\eta')\,dx,
$$

---

[1] See Bliss, <u>The problem of Lagrange</u>, loc.cit. pp. 694 - 695.

[2] Morse, <u>Sufficient conditions in the problem of</u>

where x_1, x_2 still belong to the ends of E_{12}. The function $2q(\xi, \eta)$ is a quadratic form in ξ_1, ξ_2, η_{i1}, η_{i2} with coefficients uniquely determined by E_{12}. We shall not need the explicit form of $q(\xi, \eta)$ and hence will not give it here. The function $2\omega(x, \eta, \eta')$ is given by

$$2\omega(x,\eta,\eta') = P_{ik}\eta_i\eta_k + 2Q_{ik}\eta_i\eta_k' + R_{ik}\eta_i'\eta_k', \qquad (i,k=1,\cdots,n) \tag{2.3}$$

in which P_{ik}, Q_{ik}, R_{ik} denote respectively the partial derivatives $F_{y_iy_k}$, $F_{y_iy_k'}$, $F_{y_i'y_k'}$ with arguments belonging to E_{12}. We notice that P_{ik} and R_{ik} are symmetric matrices.

If E_{12} is a minimizing arc, we must have $I_2(\xi, \eta) \geqq 0$ for every set of admissible variations ξ_1, ξ_2, $\eta_i(x)$ satisfying the conditions $\Psi_\mu(\xi, \eta) = 0$. Hence the following accessory problem naturally suggests itself, namely, to minimize the expression $I_2(\xi, \eta)$ with respect to all parameters ξ_1, ξ_2 and functions $\eta_i(x)$ of class D' on x_1x_2 and satisfying the conditions

$$\Phi_\alpha(x,\eta,\eta') = 0, \quad \Psi_\mu(\xi,\eta) = 0, \quad \int_{x_1}^{x_2}(\xi_1^2+\xi_2^2+\eta_i\eta_i)\,dx = 1. \tag{2.4}$$

This problem is rather unsymmetrical due to the presence of ξ_1, ξ_2, and will cause some trouble in its discussion unless certain restrictive conditions are introduced. However, by a simple change of notation, we can put it into a more elegant and simpler form and without any loss of generality. We shall enlarge the ranges of i, k and α so that i, k = 1, ..., n + 2; α = 1, ..., m + 2. Instead of ξ_1, ξ_2 we write η_{n+1}, η_{n+2} respectively, and define

---

Lagrange with variable end points, loc. cit., p. 521.

$$(2\cdot5)\qquad \begin{aligned} &P_{ik} = Q_{ik} = R_{ik} = 0 \quad \text{for } i \text{ or } k > n, \\ &\Phi_{m+s}(x,\eta,\eta') = \eta'_{n+s} \quad \text{for } s = 1 \text{ or } 2 . \end{aligned}$$

With these agreements and obvious changes in writing $g(\xi,\eta)$ and $\Psi_\mu(\xi,\eta)$, we see that the above problem is included in the following Problem 2, in which, however, we still keep on writing i, k = 1, ···, n; α = 1, ···, m. This is a matter of convenience because m and n are arbitrary integers except m < n. Moreover by means of the hypotheses (a), (b), (c), (d) made in § 1 one readily verifies that the hypotheses (α), ···, (ϵ) to be made below are satisfied by the accessory problem for E_{12} after change of notation.

<u>Problem</u> 2. To minimize an expression J(η) of the form

$$(2\cdot6)\qquad \begin{aligned} &J(\eta) = 2g(\eta) + \int_{x_1}^{x_2} 2\omega(x,\eta,\eta')\,dx, \text{ where} \\ &2g(\eta) = A_{ik}\eta_{i1}\eta_{k1} + 2B_{ik}\eta_{i1}\eta_{k2} + C_{ik}\eta_{i2}\eta_{k2}, \\ &2\omega(x,\eta,\eta') = P_{ik}\eta_i\eta_k + 2Q_{ik}\eta_i\eta'_k + R_{ik}\eta'_i\eta'_k, \qquad (i,k = 1,\cdots,n) \end{aligned}$$

with respect to all functions $\eta_i(x)$ of class D' on a fixed interval x_1x_2 and satisfying the conditions

$$(2\cdot7)\qquad \begin{aligned} &\Phi_\alpha(x,\eta,\eta') = \varphi_{\alpha i}\eta'_i + \theta_{\alpha i}\eta_i = 0, \quad N(\eta) = \int_{x_1}^{x_2}\eta_i\eta_i\,dx = 1, \\ &\Psi_\mu(\eta) = a_{\mu i}\eta_{i1} + c_{\mu i}\eta_{i2} = 0, \quad (\alpha = 1,\cdots,m<n;\ \mu = 1,\cdots,p \le 2n). \end{aligned}$$

The hypotheses to be made on this problem are as follows:

(α) The coefficients of $2g(\eta)$ and $\Psi_\mu(\eta)$ are all constants. The coefficients of 2ω and Φ_α are functions of x of class C' on x_1x_2. The matrices A_{ik}, C_{ik}, P_{ik}, R_{ik} are all symmetric.

(β) The m by n matrix $\|\varphi_{\alpha i}\|$ has rank m at every point of x_1x_2.

(γ) The p by 2n matrix $\|a_{\mu i}, c_{\mu i}\|$ has rank p.

(δ) There exist p sets of admissible arcs $\eta_i^\nu(x)$, ($\nu = 1, \cdots, p$), i.e., solutions of class D' of the equations $\Phi_\alpha(x, \eta, \eta') = 0$, such that

$$(2\cdot8) \qquad | \Psi_\mu(\eta^\nu) | \neq 0 , \qquad (\mu, \nu = 1, \cdots, p).$$

(ϵ) The determinant

$$(2\cdot9) \qquad R = \begin{vmatrix} R_{ik} & \bar{\varphi}_{i\beta} \\ \varphi_{\alpha k} & 0 \end{vmatrix} \qquad \left(\begin{matrix} i, k = 1, \cdots, n \\ \alpha, \beta = 1, \cdots, m \end{matrix} \right)$$

with $\bar{\varphi}_{i\beta} = \varphi_{\beta i}$, is different from zero at every point of x_1x_2.

This problem has the same form as Problem 1 except the condition $N(\eta) = 1$. However we know that it can be put exactly in the form of Problem 1 by introducing another new variable. By means of the hypothesis (δ), one can readily prove that every admissible arc for this problem is normal. The following theorem is then easily deducible from Theorem 1·1.

Theorem 2·1. For every minimizing arc Γ_{12} of Problem 2 there exists a set of constants σ, c_i, d_μ and a function

$$(2\cdot10) \qquad \Omega(x, \eta, \eta', \nu) = \omega + \nu_\alpha(x)\Phi_\alpha$$

such that the equations

$$(2\cdot11) \qquad \Omega_{\eta_i'} = \int_{x_1}^{x} (\Omega_{\eta_i} - \sigma\eta_i)\,dx + c_i$$

are satisfied at every point of Γ_{12} and the transversality conditions

$$(2\cdot12) \qquad -\Omega_{\eta_i'}(x_1) + g_{\eta_{i1}} + d_\mu a_{\mu i} = 0 , \quad \Omega_{\eta_i'}(x_2) + g_{\eta_{i2}} + d_\mu c_{\mu i} = 0 ,$$

where $\Omega_{\eta_i'}(x_1)$ and $\Omega_{\eta_i'}(x_2)$ denote the values of $\Omega_{\eta_i'}$ at x_1 and x_2 respectively, are satisfied by the ends of Γ_{12}. The func-

tions $\nu_\alpha(x)$ are continuous on x_1x_2 except possibly at values of x defining corners of Γ_{12}. Furthermore the set of constants and multipliers σ, c_i, d_μ and $\nu_\alpha(x)$ are unique.

Corollaries like those to Theorem 1·1 can be immediately written down. In particular, if Γ_{12} is a minimizing arc without corners, the equations

$$(2\cdot13) \qquad \Phi_\alpha(x,\eta,\eta')=0\,, \quad (d/dx)\Omega_{\eta_i'}=\Omega_{\eta_i}-\sigma\eta_i$$

must be satisfied at every point of Γ_{12}.

3. The boundary value problem and its canonical form. The boundary value problem associated with Problem 2 is to find for what values of σ the equations (2·13) can have a non-identically vanishing solution $\eta_i(x)$, $\nu_\alpha(x)$, which together with some constants d_μ will satisfy the transversality conditions (2·12) and the end conditions $\Psi_\mu = 0$. But, both theoretically and practically, it is found more convenient to transform (2·13) to a canonical form and eliminate d_μ from (2·12). For this purpose we introduce the new variables

$$(3\cdot1) \qquad \zeta_i=\Omega_{\eta_i'}=Q_{ki}\eta_k+R_{ik}\eta_k'+\nu_\alpha\varphi_{\alpha i}\,.$$

Since the determinant R in (2·9) is everywhere different from zero on x_1x_2, we can always solve the equations

$$(3\cdot2) \qquad \begin{aligned}\zeta_k &= Q_{ik}\eta_i+R_{ki}\eta_i'+\nu_\alpha\varphi_{\alpha k}\,,\\ \Phi_\alpha &= \varphi_{\alpha i}\eta_i'+\theta_{\alpha i}\eta_i=0\,,\end{aligned}$$

for the functions η_i' and ν_α. If we let

$$(3\cdot3) \qquad R^{-1}=\begin{vmatrix} M_{ik} & \bar{H}_{i\beta} \\ H_{\alpha k} & K_{\alpha\beta}\end{vmatrix}\,, \qquad \begin{pmatrix} i,k=1,\cdots,n \\ \alpha,\beta=1,\cdots,m\end{pmatrix}$$

with M_{ik}, $K_{\alpha\beta}$ symmetric and $H_{i\beta} = H_{\beta i}$, denote the reciprocal of R, the solution will be

$$(3\cdot4)\qquad \begin{aligned} \eta_i' &= -(M_{il}Q_{kl} + H_{\beta i}\theta_{\beta k})\eta_k + M_{ik}\zeta_k ,\\ \nu_\alpha &= -(H_{\alpha l}Q_{kl} + K_{\alpha\beta}\theta_{\beta k})\eta_k + H_{\alpha k}\zeta_k , \end{aligned} \qquad \begin{pmatrix} i,k,l = 1,\cdots,n \\ \alpha,\beta = 1,\cdots,m \end{pmatrix}$$

and hence $\Omega_{\eta_i} = P_{ik}\eta_k + Q_{ik}\eta_k' + \nu_\alpha\theta_{\alpha i}$ will be given by

$$(3\cdot5)\qquad \begin{aligned} \Omega_{\eta_i} &= (P_{ik} - Q_{ij}M_{jl}Q_{kl} - Q_{il}H_{\beta l}\theta_{\beta k} - \theta_{\alpha i}H_{\alpha l}Q_{kl} - \theta_{\alpha i}K_{\alpha\beta}\theta_{\beta k})\eta_k \\ &\quad + (Q_{il}M_{lk} + \theta_{\alpha i}H_{\alpha k})\zeta_k , \quad (i,j,k,l = 1,\cdots,n ;\ \alpha,\beta = 1,\cdots,m). \end{aligned}$$

From these relations it follows that (2·13) is equivalent to the system

$$(3\cdot6)\qquad \begin{aligned} \eta_i' &= L_{ik}\eta_k + M_{ik}\zeta_k ,\\ \zeta_i' &= N_{ik}\eta_k - L_{ki}\zeta_k - \sigma\eta_i , \end{aligned}$$

where L_{ik} is the coefficient of η_k in the first equation of (3·4) and N_{ik} is the coefficient of η_k in (3·5). From the definitions of M_{ik}, N_{ik}, the symmetric property of P_{ik}, R_{ik}, and the hypothesis that $\|\varphi_{\alpha i}\|$ is everywhere of rank m on x_1x_2, we can easily verify the following essential property of (3·6).

P_1) <u>Both the matrices</u> $\|M_{ik}\|$ <u>and</u> $\|N_{ik}\|$ <u>are symmetric and the former has rank</u> n - m <u>at every point of</u> x_1x_2.

Next write the transversality conditions (2·12) in the form

$$(3\cdot7)\qquad \begin{aligned} -\zeta_{i1} + A_{ik}\eta_{k1} + B_{ik}\eta_{k2} + d_\mu a_{\mu i} &= 0 ,\\ \zeta_{i2} + B_{ki}\eta_{k1} + C_{ik}\eta_{k2} + d_\mu c_{\mu i} &= 0 , \end{aligned} \qquad \begin{pmatrix} i,k = 1,\cdots,n \\ \mu = 1,\cdots,p \end{pmatrix}.$$

Since the p by 2n-dimensional matrix $\|a_{\mu i}, c_{\mu i}\|$ is by hypothesis of rank p, there exist exactly 2n - p linearly independent sets of constants (b_{ri}, d_{ri}), (r = p+1, ···, 2n), each of which is

orthogonal to every row of $\| a_{\mu i}, c_{\mu i} \|$. Multiplying (3·7) successively by these sets of constants, we shall obtain 2n - p linearly independent conditions free from d_μ. Adjoining $\Psi_\mu = 0$ to these conditions, we have the following system of 2n linearly independent conditions on the end values of η_i and ζ_i :

$$(3\cdot 8) \qquad \begin{aligned} S_\mu(\eta,\zeta) &= a_{\mu i}\eta_{i1} - 0\cdot\zeta_{i1} + c_{\mu i}\eta_{i2} + 0\cdot\zeta_{i2} = 0\,, \\ S_r(\eta,\zeta) &= a_{ri}\eta_{i1} - b_{ri}\zeta_{i1} + c_{ri}\eta_{i2} + d_{ri}\zeta_{i2} = 0\,, \end{aligned} \qquad \left(\begin{matrix} \mu = 1, \cdots, p \\ r = p+1, \cdots, 2n \end{matrix}\right)$$

where $S_\mu = 0$ is another way of writing $\Psi_\mu = 0$ and

$$(3\cdot 9) \qquad a_{ri} = b_{rk}A_{ik} + d_{rk}B_{ik}\,, \quad c_{ri} = b_{rk}B_{ki} + d_{rk}C_{ik}\,.$$

Hereafter we shall write the conditions (3·8) as $S_\rho(\eta,\zeta) = 0$, ($\rho = 1, \cdots, 2n$). From the definition of (b_{ri}, d_{ri}), the expressions (3·9), and the symmetric property of A_{ik}, C_{ik}, we have the following essential property of $S_\rho = 0$.

P$_2$) <u>The rows of $\| a_{\rho i}, c_{\rho i} \|$ and $\| b_{\rho i}, d_{\rho i} \|$ are mutually conjugate</u>, i.e.

$$(3\cdot 10) \qquad a_{\rho i}b_{\tau i} + c_{\rho i}d_{\tau i} = a_{\tau i}b_{\rho i} + c_{\tau i}d_{\rho i}\,, \qquad (\rho,\tau = 1, \cdots, 2n).$$

The equations (3·6) and the conditions (3·8) constitute a canonical form of our boundary value problem with the properties P$_1$) and P$_2$). In the following section we shall show that these are the only essential properties of the problem. Before leaving this section we shall state the following lemma which is easily deducible from the above transformation and will be useful in the sequel.

<u>Lemma</u> 3·1. C$_1$) <u>For every solution (η,ζ) of the conditions</u> $S_\rho(\eta,\zeta) = 0$, <u>there exist constants</u> d_μ <u>such that</u>

$$(3\cdot 11) \qquad d_\mu \Psi_\mu(\bar\eta) + g(\eta,\bar\eta) + \zeta_i\bar\eta_i \Big|_1^2 = 0$$

is an identity in $(\bar\eta)$, where $q(\eta,\bar\eta)$ is the polarized form of $2q(\eta)$ with respect to $(\bar\eta)$. C_2) For every solution (η,ζ) of the conditions $S_\rho(\eta,\zeta)=0$, we have

$$2q(\eta)+\zeta_i\eta_i\Big|_1^2=0\ . \tag{3·12}$$

C_3) For every two solutions (η,ζ) and $(\bar\eta,\bar\zeta)$ of the conditions $S_\rho(\eta,\zeta)=0$, we have

$$(\zeta_i\bar\eta_i-\bar\zeta_i\eta_i)\Big|_1^2=0\ . \tag{3·13}$$

Their proofs are immediate. For C_1) simply says that, whenever (η,ζ) satisfies the conditions $S_\rho=0$, there must be constants d_μ which together with (η,ζ) satisfy the conditions (3·7). (3·12) is obtained from (3·11) by letting $(\bar\eta)=(\eta)$ and noting that $\Psi_\mu(\eta)=0$. (3·13) is obtained by subtracting (3·11) and its analogue.

4. The equivalence of the boundary value problem and a minimizing problem. In the preceding section we have shown that Problem 2 leads to a boundary value problem consisting of (3·6), (3·8) and having properties P_1), P_2). In this section we wish to show that, conversely, any such boundary value problem is the problem arising from a minimizing problem of the same form as Problem 2.

First let

$$S_\rho(\eta,\zeta)=a_{\rho i}\eta_{i1}-b_{\rho i}\zeta_{i1}+c_{\rho i}\eta_{i2}+d_{\rho i}\zeta_{i2}=0\ , \qquad (\rho=1,\cdots,2n)$$

be any $2n$ linearly independent conditions having the conjugate property P_2). Let the rank of $\| b_{\rho i},\ d_{\rho i}\|$ be $2n-p$, $(0\le p\le 2n)$. By replacing the conditions $S_\rho=0$ by their linear combinations, we may suppose that they have been reduced to a form in which

$b_{\mu i} = d_{\mu i} = 0$, ($\mu = 1, \cdots, p$; $i = 1, \cdots, n$), and the matrices $\| a_{\rho i}, c_{\rho i} \|$ and $\| b_{\rho i}, d_{\rho i} \|$, after certain same permutation of columns, will become

$$(4\cdot1) \qquad \left\| \begin{matrix} \alpha_{\mu\nu} & \gamma_{\mu s} \\ \alpha_{\lambda\nu} & \gamma_{\lambda s} \end{matrix} \right\| \quad \text{and.} \quad \left\| \begin{matrix} 0 & 0 \\ \beta_{\lambda\nu} & \delta_{\lambda s} \end{matrix} \right\| , \qquad \left(\begin{matrix} \mu, \nu = 1, \cdots, p \\ \lambda, s = p+1, \cdots, 2n \end{matrix} \right)$$

respectively, where $\delta_{\lambda s}$ is the Kronecker δ . It can be seen without difficulty that the matrices in (4·1) will still have the conjugate property P_2). We then define $\Psi_\mu(\eta) = 0$ as the first p conditions $S_\mu(\eta, \zeta) = 0$, and prove that there exists a quadratic form $2q(\eta)$ in η_{i1} , η_{i2} with constant coefficients A_{ik}, $2B_{ik}$, C_{ik} such that the last $2n - p$ conditions $S_\lambda(\eta, \zeta)$ $= 0$ are the result of eliminating d_μ from the conditions

$$(4\cdot2) \qquad \begin{aligned} \zeta_{i1} &= A_{ik}\eta_{k1} + B_{ik}\eta_{k2} + d_\mu a_{\mu i} , \\ -\zeta_{i2} &= B_{ki}\eta_{k1} + C_{ik}\eta_{k2} + d_\mu c_{\mu i} . \end{aligned}$$

All we have to show is that the general solution ζ_{i1} , $-\zeta_{i2}$ of the equations

$$(4\cdot3) \qquad b_{\lambda i}\zeta_{i1} - d_{\lambda i}\zeta_{i2} = a_{\lambda i}\eta_{i1} + c_{\lambda i}\eta_{i2} , \qquad (\lambda = p+1, \cdots, 2n)$$

is of the form (4·2) with $A_{ik} = A_{ki}$ and $C_{ik} = C_{ki}$. To do this, let y_ρ and z_ρ denote respectively the sets (η_{i1} , η_{i2}) and (ζ_{i1} , $-\zeta_{i2}$) permuted exactly in the same way and so that, when we do the same permutation with the columns of $\| a_{\rho i}, c_{\rho i} \|$ and $\| b_{\rho i}, d_{\rho i} \|$, they will become respectively the matrices in (4·1). With these notations the equations (4·3) then become

$$(4\cdot4) \qquad \beta_{\lambda\nu} z_\nu + \delta_{\lambda s} z_s = \alpha_{\lambda\nu} y_\nu + \gamma_{\lambda s} y_s , \qquad \left(\begin{matrix} \nu = 1, \cdots, p \\ \lambda, s = p+1, \cdots, 2n \end{matrix} \right).$$

Letting $\epsilon_{\lambda s} = \gamma_{\lambda s} - \alpha_{s\nu}\beta_{\lambda\nu}$ and remembering that $\delta_{\lambda s}$ is the

Kronecker δ, we see that the general solution of (4•4) in z_μ and z_r is

$$
(4\cdot5)\qquad \begin{aligned} z_\mu &= 0\cdot y_\nu + \alpha_{s\mu} y_s + c_\nu \alpha_{\nu\mu}\,, \\ z_\lambda &= \alpha_{\lambda\nu} y_\nu + \epsilon_{\lambda s} y_s + c_\nu \gamma_{\nu s}\,. \end{aligned} \qquad \begin{pmatrix} \mu,\nu = 1,\cdots,p \\ \lambda, s = p+1,\cdots,2n \end{pmatrix}
$$

with arbitrary constants c_ν. This is evident because (4•5) with $c_\nu = 0$ is a particular solution of (4•4) by actual substitution and each row of $\| \alpha_{\mu\nu}, \gamma_{\mu s} \|$ is orthogonal to every row of $\| \beta_{\lambda\nu}, \delta_{\lambda s} \|$ by the conjugate property of the matrices in (4•1). The latter property also shows that $\epsilon_{\lambda s} = \epsilon_{s\lambda}$. The equations (4•2) then follow from (4•5) by return to our original notation since the symmetry of the matrix formed by the coefficients of y_ν, y_s in (4•5) implies that of the corresponding matrix in (4•2).

Next consider the differential equations (3•6). We are going to show that, given any functions L_{ik} and M_{ik}, N_{ik} having the property P_1), there exist a quadratic form $2\omega(x,\eta,\eta')$ in η_i, η'_i with coefficients P_{ik}, $2Q_{ik}$, R_{ik} and a set of differential equations $\Phi_\alpha(x,\eta,\eta') = \varphi_{\alpha i}\eta'_i + \theta_{\alpha i}\eta_i = 0$ such that the transformations in § 3 will lead to equations (3•6) with the given functions L_{ik}, M_{ik}, N_{ik} as coefficients. To this effect we proceed as follows: Since M_{ik} is everywhere of rank $n - m$ on x_1x_2, there exist[1] m sets of functions $H_{\alpha k}$, ($\alpha = 1, \cdots, m$), having any required continuity properties, such that $\left\| \begin{matrix} M_{ik} \\ H_{\alpha k} \end{matrix} \right\|$ is everywhere of rank n on x_1x_2. Let $\bar{H}_{i\beta} = H_{\beta i}$ and $K_{\alpha\beta} = K_{\beta\alpha}$, ($\alpha, \beta = 1, \cdots, m$) be any arbitrary functions. I say the symmetric determinant

---

[1] See Bliss, The problem of Mayer, loc.cit. p. 312. The theorem given there is a little different from the above state-

$$R^{-1} = \begin{vmatrix} M_{ik} & \bar{H}_{i\beta} \\ H_{\alpha k} & K_{\alpha\beta} \end{vmatrix} \qquad \begin{pmatrix} i,k = 1, \cdots, n \\ \alpha,\beta = 1, \cdots, m \end{pmatrix}$$

is everywhere different from zero on $x_1 x_2$ and the cofactor of any element $K_{\alpha\beta}$ is everywhere zero. Indeed let x have any fixed value and, for the sake of convenience, let the determinant of order $n - m$ in the upper left hand corner of $|M_{ik}|$ be different from zero at x. Divide the elements of R^{-1} into blocks as indicated by the following diagram

$$R^{-1} = \begin{array}{c|c|c|c} n-m & m & m & \\ \hline \mathrm{I} & \mathrm{III} & \mathrm{VI} & n-m \\ \hline \mathrm{II} & \mathrm{IV} & \mathrm{VIII} & m \\ \hline \mathrm{V} & \mathrm{VII} & \mathrm{IX} & m \end{array}$$

By subtracting suitable linear combinations of the first $n - m$ rows and columns of R^{-1} from its remaining rows and columns respectively, it can be brought about that the elements in the blocks II to VI all reduce to zero. Let the blocks VII and VIII become VII' and VIII'. By our choice of the functions $H_{\alpha k}$, neither of the determinants |VII'| and |VIII'| can be zero. It then follows that R^{-1} is different from zero at x. If we do the analogous process with the cofactor of any element $K_{\alpha\beta}$, we shall arrive at blocks VII" and VIII" of dimensions $m - 1$ by m and m by $m - 1$ respectively, and the cofactor is seen to be zero at x. Consequently the reciprocal R of R^{-1} exists everywhere on $x_1 x_2$ and is of the form given in (2·9). Now let R_{ik} and $\varphi_{\alpha i}$ be elements of this reciprocal R as they stand in (2·9) and define

$$\theta_{\alpha i} = -\varphi_{\alpha k} L_{ki}, \quad Q_{ki} = -R_{il} L_{lk}, \quad P_{\cdot k} = N_{\cdot k} - Q_{\cdot l} L_{lk} \cdot$$

---

ment but essentially the same.

We see that R_{ik} and P_{ik} are both symmetric matrices. By taking these functions P_{ik}, $2Q_{ik}$, R_{ik} and $\varphi_{\alpha i}$, $\theta_{\alpha i}$ as the coefficients in $2\omega(x,\eta,\eta')$ and $\Phi_\alpha(x,\eta,\eta')$ respectively and substituting them in the formulas of § 3, it will come out that the coefficients L_{ik}, M_{ik}, N_{ik} have exactly the given values.

5. <u>An analogue of the Jacobi condition in terms of characteristic numbers.</u> Write the equations (3·6) and the boundary conditions (3·8) together in the system

$$(5\cdot1) \qquad \begin{aligned} \eta_i' &= L_{ik}\eta_k + M_{ik}\zeta_k , \\ \zeta_i' &= N_{ik}\eta_k - L_{ki}\zeta_k - \sigma\eta_i , \end{aligned} \qquad S_\rho(\eta,\zeta)=0 , \qquad \left(\begin{matrix} i,k=1,\cdots,n \\ \rho=1,\cdots,2n \end{matrix}\right).$$

Any value of σ, for which (5·1) has a continuous and non-identically vanishing solution in η_i, ζ_i, is called a <u>characteristic number</u>, and the solution a <u>characteristic solution</u>. With these definitions we have the following

<u>Lemma 5·1. The normality condition (δ) in Problem 2 is a necessary and sufficient condition for the system (5·1) to have no characteristic solution with $\eta_i(x)$ all identically zero on x_1x_2.</u>

<u>It is necessary.</u> For if the determinant in (2·8) is zero for every p sets of admissible arcs of Problem 2, then there exist constants d_μ not all zero such that $d_\mu \Psi_\mu(\bar\eta) = 0$ for every admissible arc $\bar\eta_i$ of Problem 2. On the other hand we can always select n - m functions $\Phi_r(x,\eta,\eta')$, (r = m+1, ⋯, n) of the same form as $\Phi_\alpha(x,\eta,\eta')$ and n continuous multipliers $\nu_l(x)$, (l = 1, ⋯, n) such that[1]

[1] Bliss, <u>The problem of Mayer</u>, loc.cit. p. 312, <u>The problem of Lagrange</u>, loc.cit., p. 680.

$$|\varphi_{\ell i}| \neq 0 \quad \text{and} \quad \nu_\ell \varphi_{\ell i} = \int_{x_1}^{x} \nu_\ell \theta_{\ell i}\, dx + c_i$$

are satisfied at every point of x_1x_2 for arbitrarily chosen constants c_i. Letting $\bar{\eta}_i$ be any admissible arc of Problem 2 and $\Phi_r(x,\eta,\eta') = \tau_r$, we then have

$$\int_{x_1}^{x_2} \nu_r \tau_r\, dx = \int_{x_1}^{x_2} (\nu_\ell \varphi_{\ell i} \bar{\eta}_i' + \nu_\ell \theta_{\ell i} \bar{\eta}_i)\, dx = \nu_\ell \varphi_{\ell i} \bar{\eta}_i \Big|_1^2 .$$

By this and the fact that $d_\mu \Psi_\mu(\bar{\eta}) = 0$ we have

$$\int_{x_1}^{x_2} \nu_r \tau_r\, dx = (-c_i + d_\mu a_{\mu i}) \bar{\eta}_{i1} + (\nu_\ell \varphi_{\ell i}|^2 + d_\mu c_{\mu i}) \bar{\eta}_{i2}$$

which holds for every admissible arc $\bar{\eta}_i$ of Problem 2. On choosing $c_i = d_\mu a_{\mu i}$, the usual well-known argument then shows that the coefficients of $\bar{\eta}_{i2}$ are all zero and the multipliers ν_r all vanish identically on x_1x_2. This being so let ν_r be the remaining multipliers thus determined, $\eta_i \equiv 0$, and $\Omega(x,\eta,\eta',\nu)$ be defined by (2·10) with these arguments. It then follows easily that $\eta_i \equiv 0$, $\zeta_i = \Omega_{\eta_i'}$ form a continuous solution of (5·1) with $\zeta_{i1} = c_i = d_\mu a_{\mu i}$, $-\zeta_{i2} = d_\mu c_{\mu i}$ not all zero.

<u>It is sufficient</u>. Suppose that (5·1) has a characteristic solution η_i, ζ_i with η_i, and consequently $M_{ik}\zeta_k$, all identically zero on x_1x_2. Since (5·1) is a homogeneous system, the end values ζ_{i1}, ζ_{i2} cannot be all zero. By means of this fact and the formula (3·11) it follows that there exist constants d_μ not all zero such that

$$d_\mu \Psi_\mu(\bar{\eta}) + \zeta_i \bar{\eta}_i \Big|_1^2 = 0 \tag{5·2}$$

is an identity in $\bar{\eta}_{i1}$, $\bar{\eta}_{i2}$. Now let $\bar{\eta}_i$ be an admissible arc for Problem 2. By the transformations in § 3, $\bar{\eta}_i$ is a continuous solution of

$$\bar{\eta}_i' = L_{ik}\bar{\eta}_k + M_{ik}\bar{\zeta}_k \tag{5·3}$$

with some functions $\bar{\zeta}_i$. By multiplying (5·3) by ζ_i, the second set of the differential equations in (5·1) by $\bar{\eta}_i$, adding the results, and integrating from x_1 to x_2, it follows that $\zeta_i \bar{\eta}_i \Big|_1^2 = 0$. This and (5·2) show that $d_\mu \Psi_\mu(\bar{\eta}) = 0$ for every admissible arc $\bar{\eta}_i$ of Problem 2. Since d_μ are not all zero, this can hold only if the determinant in (2·8) is zero for every p sets of admissible arcs of Problem 2, contradictory to hypothesis.

Since the normality condition (δ) in Problem 2 is assumed throughout this paper, the above lemma shows that we can always norm the functions $\eta_i(x)$ of a characteristic solution of (5·1) so that they satisfy the norming condition $N(\eta) = 1$ in (2·7). Such a solution will be called a <u>normed</u> solution. Many other properties of the system (5·1) follow from its properties P_1), P_2) and the preceding lemma. In the following chapter we shall make a comprehensive study of this system and in Chapter III we shall see that a complete solution of it leads to a complete solution of Problem 2. In the remainder of this section we shall derive another lemma which will be useful later and from which an analogue of the Jacobi condition follows immediately.

For convenience let us say that a function f(x) is of <u>class</u> D on x_1x_2[1] if it has at most a finite member of ordinary discontinuities; i.e. it is continuous on x_1x_2 or the interval x_1x_2 can be divided into several subintervals such that it is

---

[1] This definition will be used throughout the sequel. It is new because in Bolza's <u>Vorlesungen über Variationsrechnung</u>, p. 63, only functions of class $D^{(p)}$ ($p \geq 1$) have been defined.

continuous on each of them. Having this definition let η_i be any admissible arc of Problem 2, $\Omega(x,\eta,\eta',\nu)$ be defined by (2·10) with any multipliers ν_α of class D, and $\zeta_i = \Omega_{\eta_i'}$. We see that the functions ζ_i are also of class D and the transformations in § 3 will lead to

(5·4) $\quad \eta_i' = L_{ik}\eta_k + M_{ik}\zeta_k, \qquad$ (5·5) $\quad \Omega_{\eta_i} = N_{ik}\eta_k - L_{ki}\zeta_k.$

Conversely any solution η_i, ζ_i of (5·4) with η_i continuous and ζ_i of class D determines an admissible arc for Problem 2. Consequently we may regard the totality of such solutions as the totality of admissible arcs for Problem 2. If the given admissible arc η_i is of class D" and the multipliers are taken to be of class C' or D' at pleasure, then the functions ζ_i are of class D' except possibly having a finite number of ordinary discontinuities. Hence η_i, ζ_i will be a solution

$$(5\cdot6) \qquad \begin{aligned} \eta_i' &= L_{ik}\eta_k + M_{ik}\zeta_k, \\ \zeta_i' &= N_{ik}\eta_k - L_{ki}\zeta_k - g_i \end{aligned}$$

with some functions g_i of class D. With these facts we have the following

<u>Lemma</u> 5·2. C_1) <u>For any admissible arc</u> η_i, ζ_i <u>of Problem</u> 2, <u>the value of the expression</u> $J(\eta)$ <u>in</u> (2·6) <u>is</u>

$$(5\cdot7) \qquad J(\eta) = 2q(\eta) + \int_{x_1}^{x_2} (N_{ik}\eta_i\eta_k + M_{ik}\zeta_i\zeta_k)\,dx.$$

C_2) <u>If the functions</u> η_i, ζ_i <u>are a continuous solution of</u> (5·6) <u>corresponding to functions</u> g_i <u>of class</u> D, <u>then</u>

$$(5\cdot8) \qquad J(\eta) = 2q(\eta) + \zeta_i\eta_i\Big|_1^2 + \int_{x_1}^{x_2} g_i\eta_i\,dx.$$

C_3) <u>If the preceding solution</u> η_i, ζ_i <u>also satisfies the boun-</u>

dary conditions $S_\rho(\eta,\zeta)=0$ in (5·1), then

$$J(\eta)=\int_{x_1}^{x_2} g_i\eta_i\,dx. \tag{5·9}$$

C_4) If the functions η_i, ζ_i are a normed characteristic solution of (5·1), then

$$J(\eta)=\sigma\int_{x_1}^{x_2}\eta_i\eta_i\,dx=\sigma. \tag{5·10}$$

Their proofs are immediate. Because we have

$$J(\eta)=2g(\eta)+\int_{x_1}^{x_2}2\Omega\,dx=2g(\eta)+\int_{x_1}^{x_2}(\Omega_{\eta_i}\eta_i+\Omega_{\eta_i'}\eta_i'+\Omega_{\nu_\alpha}\nu_\alpha)dx.$$

The relation (5·7) then follows from this by writing ζ_i for $\Omega_{\eta_i'}$, substituting the values of η_i' and Ω_{η_i} from (5·4), (5·5), and noting that $\Omega_{\nu_\alpha}=\Phi_\alpha=0$. If η_i, ζ_i form a continuous solution of (5·6), then (5·8) follows from (5·7) by multiplying the first line of (5·6) by ζ_i, the second by η_i, adding the results, and integrating from x_1 to x_2. (5·9) follows from (5·8) and (3·12). Finally (5·10) follows from (5·9).

The following corollary gives an analogue of the Jacobi condition.

Corollary. If E_{12} is a normal non-singular minimizing arc of Problem 1 and is without corners, then its accessory boundary value problem can have no negative characteristic numbers.

CHAPTER II.

A STUDY OF THE BOUNDARY VALUE SYSTEM.

With the properties P_1) and P_2) of the system (5·1), we can show that it is self-adjoint but, in general, not definitely self-adjoint according to the definition of Bliss,[1] unless the determinant $|M_{ik}|$ is everywhere different from zero on x_1x_2 and hence $m = 0$. However most of the theorems as given by Bliss for definitely self-adjoint systems, with some modifications, can be proved for this system by essentially the same process. In the following section we shall first give a few lemmas which are easily deducible from the calculus of variations property of the system, and thus make our whole argument independent of any previous knowledge of the theory of self-adjoint systems. It is also to be noted that, as a consequence of the hypotheses in Problem 2, the coefficients L_{ik}, M_{ik}, N_{ik} in (5·1) are all of class C'. But throughout the following discussions only the property of continuity is involved.

6. Preliminary lemmas. Let $\eta_i^\tau(x,\sigma)$, $\zeta_i^\tau(x,\sigma)$, ($\tau = 1, \cdots, 2n$) be 2n linearly independent solutions[2] of the differential equations in (5·1) and $D(\sigma)$ denote the determinant

$$D(\sigma) = |S_\rho(\eta^\tau, \zeta^\tau)| \tag{6·1}$$

First it is obvious that any solution of (5·1) must be a linear combination $\eta_i = c_\tau \eta_i^\tau$, $\zeta_i = c_\tau \zeta_i^\tau$ of the functions η_i^τ, ζ_i^τ

---

[1] See Bliss, A boundary value problem, loc.cit., pp. 568 - 569.

[2] All solutions are assumed to be continuous unless otherwise mentioned.

with constant coefficients c_τ satisfying the conditions

$$(6\cdot2)\qquad S_\rho(\eta,\zeta) = S_\rho(c_\tau\eta^\tau, c_\tau\zeta^\tau) = c_\tau S_\rho(\eta^\tau, \zeta^\tau) = 0\,,$$

and the number of linearly independent solutions of (5•1) is exactly the number of linearly independent solutions in c_τ of these conditions. If $D(\sigma) \neq 0$, the constants c_τ must evidently be all zero. If $D(\sigma)$ is of rank $2n - r$, there exist r and only r linearly independent solutions in c_τ. Next, on using the notations

$$(6\cdot3)\qquad q_{i1}^\tau = q_{\eta_{i1}}(\eta^\tau)\,, \qquad q_{i2}^\tau = q_{\eta_{i2}}(\eta^\tau)\,, \qquad (\tau = 1, \cdots, 2n)$$

the constants c_τ together with some constants d_μ are also determined by the system of equations

$$(6\cdot4)\qquad \begin{aligned} c_\tau \Psi_\mu(\eta^\tau) &+ d_\mu \cdot 0 = 0\,, \\ c_\tau(q_{i1}^\tau - \zeta_{i1}^\tau) &+ d_\mu a_{\mu i} = 0\,, \\ c_\tau(q_{i2}^\tau + \zeta_{i2}^\tau) &+ d_\mu c_{\mu i} = 0\,, \end{aligned} \qquad \begin{pmatrix} i = 1, \cdots, n \\ \mu = 1, \cdots, p \\ \tau = 1, \cdots, 2n \end{pmatrix}.$$

This is true because this system requires the functions $c_\tau\eta_i^\tau$, $c_\tau\zeta_i^\tau$ to satisfy the end conditions $\Psi_\mu(\eta) = 0$ and the transversality conditions (3•7) which are equivalent to the boundary conditions $S_\rho(\eta,\zeta) = 0$. Furthermore since the matrix $\| a_{\mu i}, c_{\mu i} \|$ is assumed to have rank p, the system (6•4) can have no non-vanishing solution with c_τ all zero. Hence, if $D(\sigma)$ is of rank $2n - r$, the determinant

$$(6\cdot5)\qquad \Delta(\sigma) = \begin{vmatrix} \Psi_\mu(\eta^\tau) & 0 \\ q_{i1}^\tau - \zeta_{i1}^\tau & a_{\mu i} \\ q_{i2}^\tau + \zeta_{i2}^\tau & c_{\mu i} \end{vmatrix}\,, \qquad \begin{pmatrix} i = 1, \cdots, n \\ \mu = 1, \cdots, p \\ \tau = 1, \cdots, 2n \end{pmatrix}$$

for the same value of σ, must be of rank $p + 2n - r$, and

conversely. This proves

Lemma 6·1. Corresponding to a number σ, if $D(\sigma)$ is different from zero, the system (5·1) has no characteristic solution. If $D(\sigma)$ is of rank $2n - r$, the system has r and only r linearly independent solutions. Moreover, if the determinant $D(\sigma)$ is of rank $2n - r$, the determinant $\Delta(\sigma)$ is of rank $p + 2n - r$, and conversely. Their zeros therefore give the characteristic numbers of the system as defined in § 5.

Both the determinants $D(\sigma)$ and $\Delta(\sigma)$ may be called characteristic determinants of (5·1). Since the right members of the differential equations in (5·1) are analytic in σ in the whole complex σ-plane, it follows that the solutions $\eta_i^{\tau}(x,\sigma)$, $\zeta_i^{\tau}(x,\sigma)$ as well as the above determinants are also analytic in σ in the same region. Hence all of them can be expanded as permanently convergent power series in σ. Next we have the following

Lemma 6·2. Let g_i and $\bar{g}_i$ be any sets of functions of class D, and η_i, ζ_i and $\bar{\eta}_i$, $\bar{\zeta}_i$ be respectively continuous solutions of the differential equations in the following systems

$$(6\cdot6)\qquad \begin{aligned} \eta_i' &= L_{ik}\eta_k + M_{ik}\zeta_k, \\ \zeta_i' &= N_{ik}\eta_k - L_{ki}\zeta_k - g_i, \\ S_\rho(\eta,\zeta) &= 0, \end{aligned} \qquad (6\cdot7)\qquad \begin{aligned} \bar{\eta}_i' &= L_{ik}\bar{\eta}_k + M_{ik}\bar{\zeta}_k, \\ \bar{\zeta}_i' &= N_{ik}\bar{\eta}_k - L_{ki}\bar{\zeta}_k - \bar{g}_i, \\ S_\rho(\bar{\eta},\bar{\zeta}) &= 0. \end{aligned}$$

We have first the formula

$$(6\cdot8)\qquad \int_{x_1}^{x_2} (\bar{\eta}_i g_i - \eta_i \bar{g}_i)\,dx = (\bar{\zeta}_i\eta_i - \zeta_i\bar{\eta}_i)\Big|_1^2 .$$

If the above solutions also satisfy the boundary conditions $S_\rho = 0$, then

$$\int_{x_1}^{x_2} (\bar{\eta}_i g_i - \eta_i \bar{g}_i)\, dx = 0 . \tag{6·9}$$

Its proof follows from the fact that M_{ik} and N_{ik} are symmetric matrices. For multiplying the differential equations in (6·6) by $\bar{\zeta}_i$, $-\bar{\eta}_i$, those in (6·7) by $-\zeta_i$, η_i, adding the results, integrating from x_1 to x_2, and remembering the above fact, we have the formula (6·8) as desired. If the boundary conditions $S_\rho = 0$ are also satisfied, then (6·9) follows from (6·8) and (3·13).

Having the above two lemmas we can prove the following

Lemma 6·3. For any continuous functions g_i, the system (6·6) can have continuous solutions if and only if the condition

$$\int_{x_1}^{x_2} \bar{\eta}_i g_i \, dx = 0 \tag{6·10}$$

is satisfied for the functions $\bar{\eta}_i$ of every continuous solution of (6·7) with $\bar{g}_i \equiv 0$.

The necessity of the condition follows from (6·9) with $\bar{g}_i \equiv 0$. To prove its sufficiency, let η_i°, ζ_i° be a particular solution of the differential equations in (6·6) and η_i^τ, ζ_i^τ, ($\tau = 1, \cdots, 2n$) be $2n$ linearly independent solutions of the same equations with $g_i \equiv 0$. The general solution of the differential equations in (6·6) is then of the form $\eta_i = \eta_i^\circ + c_\tau \eta_i^\tau$, $\zeta_i = \zeta_i^\circ + c_\tau \zeta_i^\tau$ with arbitrary constants c_τ. By the notations in (6·3), (6·4) and the same argument as used there, this solution will be a solution of the system (6·6) itself, if and only if the constants c_τ together with some constants d_μ satisfy the conditions

$$\begin{aligned} c_\tau \Psi_\mu(\eta^\tau) + d_\mu \cdot 0 &= -\Psi_\mu(\eta^\circ) , \\ c_\tau (g_{i1}^\tau - \zeta_{i1}^\tau) + d_\mu a_{\mu i} &= -(g_{i1}^\circ - \zeta_{i1}^\circ) , \end{aligned} \tag{6·11}$$

$$c_\tau(q_{i2}^\tau+\zeta_{i2}^\tau)+d_\mu c_{\mu i}=-(q_{i2}^0+\zeta_{i2}^0)\,.$$

All we need to show is that (6•11) actually has a solution in c_τ, d_μ if the conditions (6•10) are satisfied. To do this, let the matrix formed by the coefficients of c_τ, d_μ in (6•11) have rank p + 2n - r. By Lemma 6•1, the system (6•7) with $\bar{g}_1 \equiv 0$ has exactly r linearly independent solutions $\bar{\eta}_i^\nu$, $\bar{\zeta}_i^\nu$, (ν = 1, •••, r). By (3•11), for each of these solutions, there exists a set of constants $\bar{d}_\mu^\nu$ such that

$$\bar{d}_\mu^\nu \Phi_\mu(u)+q(\bar{\eta},u)+\bar{\zeta}_i^\nu u_i\Big|_1^2=0 \tag{6•12}$$

is an identity in u_{i1}, u_{i2}. The r sets of constants ($\bar{d}_\mu^\nu$, $\bar{\eta}_{i1}^\nu$, $\bar{\eta}_{i2}^\nu$), (ν = 1, •••, r) are linearly independent. For otherwise, there would exist constants χ_ν, not all zero, such that $\chi_\nu \bar{d}_\mu^\nu=\chi_\nu\bar{\eta}_{i1}^\nu=\chi_\nu\bar{\eta}_{i2}^\nu=0$. The identities (6•12) would then require $\chi_\nu\bar{\zeta}_{i1}^\nu=\chi_\nu\bar{\zeta}_{i2}^\nu=0$, which is evidently impossible because $\chi_\nu\bar{\eta}_i^\nu$, $\chi_\nu\bar{\zeta}_i^\nu$ constitute a non-vanishing solution of the homogeneous system (6•7) with $\bar{g}_1 \equiv 0$. Having proved this, we give a fixed value to ν, multiply equations (6•11) with $\bar{d}_\mu^\nu$, $\bar{\eta}_{i1}^\nu$, $\bar{\eta}_{i2}^\nu$ respectively, and add the results. The coefficients of the d_μ in the resulting equation are clearly all zero. The coefficients of c_τ and the right member with sign changed are respectively the quantities

$$\bar{d}_\mu^\nu\Phi_\mu(\eta^\tau)+q(\bar{\eta}^\nu,\eta^\tau)+\zeta_i^\tau\bar{\eta}_i^\nu\Big|_1^2 \quad \text{and} \quad \bar{d}_\mu^\nu\Phi_\mu(\eta^0)+q(\bar{\eta}^\nu,\eta^0)+\zeta_i^0\bar{\eta}_i^\nu\Big|_1^2\,.$$

But we have $\zeta_i^\tau\bar{\eta}_i^\nu\Big|_1^2=\bar{\zeta}_i^\nu\eta_i^\tau\Big|_1^2$ by (6•8) with $g_1=\bar{g}_1=0$, and $\zeta_i^0\bar{\eta}_i^\nu\Big|_1^2=\bar{\zeta}_i^\nu\eta_i^0\Big|_1^2$ by (6•8) with $\bar{g}_1=0$ and the condition (6•10). Consequently the above quantities become respectively

$$\bar{d}_\mu^\nu\Phi_\mu(\eta^\tau)+q(\bar{\eta}^\nu,\eta^\tau)+\bar{\zeta}_i^\nu\eta_i^\tau\Big|_1^2 \quad \text{and} \quad \bar{d}_\mu^\nu\Phi_\mu(\eta^0)+q(\bar{\eta}^\nu,\eta^0)+\bar{\zeta}_i^\nu\eta_i^0\Big|_1^2\,,$$

which are all zero by the identity (6·12). This shows that each of the r linearly independent sets of constants ($\bar{d}_{\mu}^{\nu}$, $\bar{\eta}_{i1}^{\nu}$, $\bar{\eta}_{i2}^{\nu}$), which is orthogonal to the coefficients of c_{τ} and d_{μ} in (6·11), is also orthogonal to its right member. Since the matrix formed by the coefficients of c_{τ}, d_{μ} in (6·11) is assumed to have rank p + 2n - r, it follows that (6·11) actually has a solution in c_{τ}, d_{μ} as was to be proved.

7. Properties of the characteristic numbers and functions. In this section we shall study the properties of the characteristic numbers and functions of the system (5·1) in order to reach the important expansion theorem of the next section. First of all we remark that all our preceding lemmas remain true if the coefficient N_{ik} in (5·1) is replaced by $N_{ik} - c\,\delta_{ik}$, c being any constant and δ_{ik} being the Kronecker δ, because the only property of N_{ik} which we have made use of is $N_{ik} = N_{ki}$. For brevity we shall use the notation

$$N_{ik}^{c} = N_{ik} - c\,\delta_{ik}\,. \tag{7·1}$$

Theorem 7·1. All roots of the characteristic determinant $D(\sigma)$ of the system (5·1) are real and the linearly independent characteristic solutions corresponding to each root may be chosen real.

Let σ be a root of $D(\sigma)$ and η_i, ζ_i be a characteristic solution of (5·1) corresponding to it. Since the coefficients occurring in (5·1) are all real, if σ were not real, its conjugate imaginary $\bar{\sigma}$ would be a root of $D(\sigma)$ and the conjugate imaginaries $\bar{\eta}_i$, $\bar{\zeta}_i$ of η_i, ζ_i would be a characteristic solution corresponding to $\bar{\sigma}$. By (6·9) with $g_i = \sigma\,\eta_i$, $\bar{g}_i = \bar{\sigma}\,\bar{\eta}_i$, we would have

$$(\sigma-\bar{\sigma})\int_{x_1}^{x_2}\eta_i\bar{\eta}_i\,dx=0\,.$$

Hence the η_i would be all identically zero on x_1x_2, contradictory to Lemma 5·1. Therefore σ must be real. Since σ is real, the real and imaginary parts of η_i, ζ_i are both solutions of (5·1). The characteristic solutions corresponding to each root σ may then be chosen real as described in the theorem.

Theorem 7·2. The index of each characteristic number σ_0, i.e., the number of linearly independent solutions corresponding to it, is equal to the multiplicity of σ_0 as a root of $D(\sigma)$.

Suppose that $D(\sigma) = |\,S_\rho[\eta^\tau(x,\sigma),\ \zeta^\tau(x,\sigma)]\,|$ has rank $2n - r$ at a particular value σ_0. By replacing the solutions $\eta_i^\tau(x,\sigma)$, $\zeta_i^\tau(x,\sigma)$ by suitably selected linear combinations of them with constant coefficients, it may be brought about that at $\sigma = \sigma_0$ the expressions $S_\rho(\eta^{\tau_1}, \zeta^{\tau_1})$, ($\tau_1 = 1, \cdots, r$) all vanish, while the matrix of elements $S_\rho(\eta^{\tau_2}, \zeta^{\tau_2})$, ($\tau_2 = r+1, \cdots, 2n$ has rank $2n - r$. All derivatives of $D(\sigma)$ of order less than r will then clearly vanish at $\sigma = \sigma_0$, and the r-th will have the value

$$D^{(r)}(\sigma_0) = \underline{|r}\,\left|\,S_\rho(\eta_\sigma^{\tau_1}, \zeta_\sigma^{\tau_1}),\ S_\rho(\eta^{\tau_2}, \zeta^{\tau_2})\right|\,,$$

where the subscript σ indicates derivatives. If this expression vanished, there would be a linear combination

$$\eta_i = c_{\tau_1}\eta_{i\sigma}^{\tau_1} + c_{\tau_2}\eta_i^{\tau_2},\quad \zeta_i = c_{\tau_1}\zeta_{i\sigma}^{\tau_1} + c_{\tau_2}\zeta_i^{\tau_2}$$

for which all the numbers $S_\rho(\eta,\zeta)$ would vanish at $\sigma = \sigma_0$. The constants c_{τ_1} could not be all zero because the rank of the last $2n - r$ columns of $D^{(r)}(\sigma_0)$ is $2n - r$. The functions

$\bar{\eta}_i = c_{\tau_1}\eta_i^{\tau_1}$, $\bar{\zeta}_i = c_{\tau_1}\zeta_i^{\tau_1}$ would therefore not all vanish identically. For $\sigma = \sigma_0$ they would satisfy the system

$$(7\cdot2) \qquad \begin{aligned} \bar{\eta}_i' &= L_{ik}\bar{\eta}_k + M_{ik}\bar{\zeta}_k , \\ \bar{\zeta}_i' &= N_{ik}^{\sigma}\bar{\eta}_k - L_{ki}\bar{\zeta}_k , \end{aligned} \qquad S_\rho(\bar{\eta},\bar{\zeta}) = 0,$$

where N_{ik}^{σ} is given by (7·1), and the differential equations

$$(7\cdot3) \qquad \begin{aligned} \bar{\eta}_{i\sigma}' &= L_{ik}\bar{\eta}_{k\sigma} + M_{ik}\bar{\zeta}_{k\sigma} , \\ \bar{\zeta}_{i\sigma}' &= N_{ik}^{\sigma}\bar{\eta}_{k\sigma} - L_{ki}\bar{\zeta}_{k\sigma} - \bar{\eta}_i . \end{aligned}$$

The set of functions $\tilde{\eta}_i = c_{\tau_2}\eta_i^{\tau_2}$, $\tilde{\zeta}_i = c_{\tau_2}\zeta_i^{\tau_2}$ would satisfy the differential equations in (7·2), and, with the help of (7·3), it follows readily that the functions η_i , ζ_i themselves would for $\sigma = \sigma_0$ be solutions of the non-homogeneous system

$$(7\cdot4) \qquad \begin{aligned} \eta_i' &= L_{ik}\eta_k + M_{ik}\zeta_k , \\ \zeta_i' &= N_{ik}^{\sigma}\eta_k - L_{ki}\zeta_k - \bar{\eta}_i , \end{aligned} \qquad S_\rho(\eta,\zeta) = 0.$$

By applying (6·9) to the systems (7·2) and (7·4), we would have

$$\int_{x_1}^{x_2} \bar{\eta}_i \bar{\eta}_i \, dx = 0 .$$

Hence $\bar{\eta}_i$ would be all identically zero on x_1x_2 in contradiction with Lemma 5·1. Therefore the derivative $D^{(h)}(\sigma_0)$ is different from zero and σ_0 has its multiplicity equal to its index.

The following corollary is immediate.

Corollary. *The characteristic determinant $D(\sigma)$ is not identically zero in σ* .

Theorem 7·3. *If a set of functions h_i be continuous on the interval x_1x_2 and satisfy the condition*

$$(7\cdot5) \qquad \int_{x_1}^{x_2} \eta_i h_i \, dx = 0$$

<u>with the functions</u> η_i <u>of every characteristic solution of the system (5·1), then it satisfies the same condition with the functions</u> η_i <u>of every continuous solution of the system</u> (6·6) <u>corresponding to functions</u> g_1 <u>of class</u> D.

For let h_i be any functions satisfying the hypothesis of the theorem. According to Lemma 6·3 the non-homogeneous system

$$(7\cdot6) \qquad \begin{aligned} \eta_i' &= L_{ik}\eta_k + M_{ik}\zeta_k , \\ \zeta_i' &= N_{ik}^{\sigma}\eta_k - L_{ki}\zeta_k - h_i , \end{aligned} \qquad S_\rho(\eta,\zeta) = 0 ,$$

where N_{ik}^{σ} is given by (7·1), has solutions for every value of σ. When

$$D(\sigma) = \left| S_\rho[\eta^\tau(x,\sigma), \zeta^\tau(x,\sigma)] \right| \neq 0 ,$$

there is a unique solution and one verifies readily that it consists of the functions

$$\eta_i(x,\sigma) = \frac{1}{D(\sigma)} \begin{vmatrix} \eta_i^0 & \eta_i^\tau \\ S_\rho(\eta^0,\zeta^0) & S_\rho(\eta^\tau,\zeta^\tau) \end{vmatrix} , \quad \zeta_i(x,\sigma) = \frac{1}{D(\sigma)} \begin{vmatrix} \zeta_i^0 & \zeta_i^\tau \\ S_\rho(\eta^0,\zeta^0) & S_\rho(\eta^\tau,\zeta^\tau) \end{vmatrix} ,$$

where η_i^0, ζ_i^0 are a particular solution of the differential equations in (7·6).

Near a root σ_0 of multiplicity r of D(σ) the functions $\eta_i(x,\sigma)$, $\zeta_i(x,\sigma)$ are still well defined and analytic in σ. For in the first place one can add constant multiples of the last 2n columns of the above determinants to their first columns in such a way that the resulting functions η_i^0, ζ_i^0 satisfy the conditions $S_\rho(\eta^0,\zeta^0) = 0$ at $\sigma = \sigma_0$. This is possible because a solution of (7·6) always exists. In the second place, since by the preceding theorem D(σ) = $|S_\rho(\eta^\tau,\zeta^\tau)|$

has rank 2n - r at $\sigma = \sigma_0$, one can therefore replace r of the rows of the above determinants by r linear combinations of their last 2n rows which vanish at $\sigma = \sigma_0$. It is then clear that these determinants have the same factor $(\sigma - \sigma_0)^r$ as $D(\sigma)$, and hence that the functions $\eta_i(x,\sigma)$, $\zeta_i(x,\sigma)$ are analytic near σ_0 as well as near values of σ for which $D(\sigma) \neq 0$.

The functions $\eta_i(x,\sigma)$, $\zeta_i(x,\sigma)$ are therefore representable by permanently convergent power series in σ of the form

$$(7\cdot7) \qquad \eta_i(x,\sigma) = \eta_{i0}(x) + \sigma\eta_{i1}(x) + \cdots, \quad \zeta_i(x,\sigma) = \zeta_{i0}(x) + \sigma\zeta_{i1}(x) + \cdots .$$

By substituting these series in (7·6) and comparing the coefficients of σ , it is found that the coefficients $\eta_{i\mu}$, $\zeta_{i\mu}$, satisfy the systems

$$(7\cdot8) \qquad \begin{aligned} \eta'_{i\mu} &= L_{ik}\eta_{k\mu} + M_{ik}\zeta_{k\mu} , \\ \zeta'_{i\mu} &= N_{ik}\eta_{k\mu} - L_{ki}\zeta_{k\mu} - \eta_{i,\mu-1} , \end{aligned} \quad S_\rho(\eta_\mu, \zeta_\mu) = 0 , \quad (\mu = 0, 1, 2, \cdots),$$

in which it is agreed that $\eta_{i,-1} = h_i$. With the help of these systems and the relation (6·9), it follows that

$$\int_{x_1}^{x_2} \eta_{i\mu}\eta_{i,\nu-1}\,dx = \int_{x_1}^{x_2} \eta_{i\nu}\eta_{i,\mu-1}\,dx , \qquad (\mu, \nu = 0, 1, 2, \cdots).$$

The well-known Schwarz inequality

$$\left[\int_{x_1}^{x_2} \eta_{i,\mu-1}\eta_{i,\mu+1}\,dx \right]^2 \leq \int_{x_1}^{x_2} \eta_{i,\mu-1}\eta_{i,\mu-1}\,dx \int_{x_1}^{x_2} \eta_{i,\mu+1}\eta_{i,\mu+1}\,dx$$

together with the last equations shows that the constants

$$W_{-2} = \int_{x_1}^{x_2} \eta_{i,-1}\eta_{i,-1}\,dx , \quad W_\mu = \int_{x_1}^{x_2} \eta_{i0}\eta_{i\mu}\,dx , \qquad (\mu = 0, 1, 2, \cdots)$$

have the properties

$$(7\cdot9)\qquad W_{\mu+\nu}=\int_{x_1}^{x_2}\eta_{i\mu}\eta_{i\nu}\,dx\,,\quad W_{2\mu-2}W_{2\mu+2}\geq W_{2\mu}^2\,,\qquad (\mu,\nu=0,1,2,\cdots).$$

The series

$$(7\cdot10)\qquad W_0+\sigma W_1+\sigma^2W_2+\cdots\,,\quad W_0+\sigma^2W_2+\sigma^4W_4+\cdots\,,$$

the first of which is found by integrating the expression $\eta_{i0}\eta_i$ found from (7•7), both converge for every value of σ. Hence W_0 must be zero. For otherwise it would follow from the inequalities (7•9) that $W_{2\mu}\neq 0$ for every μ and $W_{2\mu}\geq W_0(W_2/W_0)^{\mu}$. The second series in (7•10) would not converge for $\sigma = (W_0/W_2)^{1/2}$, which is a contradiction. From the equality $W_0 = 0$, it follows, however, that the functions η_{i0} all vanish identically on x_1x_2. Now let η_i, ζ_i be any continuous solution of (6•6) corresponding to functions g_i of class D. By applying (6•9) to the system (6•6) and the system (7•8) with subscript $\mu = 0$, we have

$$\int_{x_1}^{x_2}\eta_i h_i\,dx=\int_{x_1}^{x_2}\eta_{i0}g_i\,dx=0$$

as was to be proved.

<u>Corollary</u>. <u>When a number σ_0 is not a characteristic number of (5•1), the unique solution of (7•6) with $\sigma = \sigma_0$ and corresponding to functions h_i satisfying the hypothesis of the theorem is a set of functions η_i, ζ_i in which η_i are all zero on x_1x_2.</u>

Its proof is immediate. For, when σ_0 is not a characteristic number of (5•1), the solution η_i, ζ_i of (7•6) with $\sigma = \sigma_0$ is evidently unique and, for functions h_i as described in the corollary, it is given by (7•7) with $\sigma = \sigma_0$. But in the above proof we have shown that $W_0 = 0$ and hence, by the

inequalities (7·9), $W_{2\mu} = 0$ for every μ. Consequently the functions η_{i0}, η_{i1}, ··· . are all identically zero on x_1x_2.

Theorem 7·4. The linearly independent characteristic solutions of the system (5·1) are denumerably infinite in number. These solutions with their corresponding characteristic numbers may be represented by the symbols $\eta_{i\nu}$, $\zeta_{i\nu}$, σ_ν, ($\nu = 1, 2, 3, \cdots$). The functions $\eta_{i\nu}$ may further be chosen orthonormal in the sense that

$$\int_{x_1}^{x_2} \eta_{i\mu}\eta_{i\nu}\,dx = \delta_{\mu\nu}\,, \qquad (\delta_{\mu\mu}=1,\ \delta_{\mu\nu}=0 \text{ if } \mu\neq\nu). \tag{7·11}$$

For let σ have any fixed value which is not a characteristic number of (5·1) and consider the system (7·6) with σ equal to this fixed value. It is evident that, for each set of functions ζ_i^τ, ($\tau = 1, \cdots, t$) of class C', there exist continuous functions η_i^τ, h_1^τ which together with ζ_i^τ satisfy the differential equations in (7·6). Let the functions ζ_i^τ be chosen so that the t sets of functions $M_{1k}\zeta_k^\tau$ are linearly independent. This is possible because not all the coefficients M_{1k} are identically zero on x_1x_2. For any constants c_τ, the linear combination $\eta_i = c_\tau\eta_i^\tau$, $\zeta_i = c_\tau\zeta_i^\tau$ is a solution of the same differential equations with $h_1 = c_\tau h_1^\tau$. If (5·1) had only a finite number of linearly independent characteristic solutions, then, for a sufficiently large value of t, constants c_τ not all zero could be chosen so that h_1 would satisfy the hypothesis in Theorem 7·3 and the functions η_i, ζ_i would satisfy the boundary conditions $S_\rho(\eta,\zeta) = 0$. By the corollary to that theorem, the functions η_i and hence $M_{1k}\zeta_k = c_\tau M_{1k}\zeta_k^\tau$ would be all identically zero on x_1x_2 in contradiction with our choice of the functions ζ_i^τ. Therefore the linearly independent

characteristic solutions of (5·1) must be infinite in number and the infinity is denumerable, because the roots of the permanently convergent power series $D(\sigma)$, which is not identically zero by the corollary to Theorem 7·2, must be denumerable and the index of each root is finite by Lemma 6·1.

Finally Lemma 5·1 shows that the integrals in (7·11) for $\mu = \nu$ are all different from zero, and the relation (6·9) shows that those integrals for $\eta_{i\mu}$, $\eta_{i\nu}$ corresponding to characteristic numbers $\sigma_\mu \neq \sigma_\nu$ are all zero because that relation gives

$$(\sigma_\mu - \sigma_\nu)\int_{x_1}^{x_2} \eta_{i\mu}\eta_{i\nu}\,dx = 0 .$$

The functions $\eta_{i\mu}$ may therefore be chosen orthonormal as described in the theorem by a process which is well known in the theory of integral equations.

8. <u>The Green's matrix and an expansion theorem.</u> In this section we shall assume that $\sigma = 0$ is not a characteristic number of (5·1), i.e., the system

$$\text{(8·1)}\qquad \begin{aligned} \eta_i' &= L_{ik}\eta_k + M_{ik}\zeta_k \\ \zeta_i' &= N_{ik}\eta_k - L_{ki}\zeta_k , \end{aligned} \qquad S_\rho(\eta,\zeta) = 0$$

is incompatible. If $D = |S_\rho(\eta^\tau, \zeta^\tau)|$ denotes the characteristic determinant of (8·1), we have $D \neq 0$. Let t be any value of the interval x_1x_2, and $u_i^\gamma(t, x)$, $v_i^\gamma(t, x)$, $(\gamma = 1, \cdots, n)$ be n sets of discontinuous solutions of the differential equations in (8·1) depending upon t in the following way: On the interval $(x_1 t)$ all the functions u_i^γ, v_i^γ are taken identically equal to zero; on the interval (tx_2) they are taken to be continuous solutions of the above differential equations with

initial values $u_i^{\chi}(t, t+0) = 0$, $v_i^{\chi}(t, t+0) = \delta_{i\ell}$, where $\delta_{i\ell}$ is the Kronecker δ. For each value of t and χ solve the equations

$$-S_\rho(u^\ell, v^\ell) + C_{\tau\ell}(t) S_\rho(\eta^\tau, \zeta^\tau) = 0$$

for the constants $c_{\tau\ell}(t)$. This is possible because $D \neq 0$. these constants define

$$(8\cdot2) \qquad G_{\ell i}(t,x) = -u_i^\ell(t,x) + C_{\tau\ell}(t)\eta_i^\tau(x), \quad F_{\ell i}(t,x) = -v_i^\ell(t,x) + C_{\tau\ell}(t)\zeta_i^\tau(x$$

It is then evident that, for each value of t and χ, the function $G_{\chi i}$ and $F_{\chi i}$ have the properties

$$(8\cdot3) \qquad \begin{gathered} F_{\ell i}(t,t-0) - F_{\ell i}(t,t+0) = \delta_{\ell i}, \\ \frac{\partial}{\partial x} G_{\ell i}(t,x) = L_{ik} G_{\ell k}(t,x) + M_{ik} F_{\ell k}(t,x), \\ \frac{\partial}{\partial x} F_{\ell i}(t,x) = N_{ik} G_{\ell k}(t,x) - L_{ki} F_{\ell k}(t,x), \\ S_\rho[G(t,x), F(t,x)] = 0 . \end{gathered}$$

Further these functions are readily seen to be bounded in (t, x), because the functions $G_{\chi i}$ are everywhere continuous and the functions $F_{\chi i}$ can have ordinary discontinuities only when $t = x$ and $\chi = i$. These functions form a part of the so-called <u>Green's matrix</u> for the system (8·1) and would be sets of characteristic solutions of the system if it were not for the discontinuities of F_{ii}. With these functions, we have the following

<u>Lemma 8·1.</u> <u>If</u> η_i, ζ_i <u>be a continuous solution of the system (6·6) corresponding to functions</u> g_i <u>of class D, the functions</u> η_i <u>are given by</u>

$$(8\cdot4) \qquad \eta_i(x) = \int_{x_1}^{x_2} G_{i\ell}(x,t) g_\ell(t)\,dt .$$

For from the systems (6•6), (8•3) and the same multiplying and adding process as used in the proof of Lemma 6•2, we see that

$$\frac{\partial}{\partial x}\left[G_{\ell i}(t,x)\zeta_i(x) - F_{\ell i}(t,x)\eta_i(x)\right] = -G_{\ell i}(t,x)g_i(x). \tag{8•8}$$

By integrating first from x_1 to $t - 0$, next from $t + 0$ to x_2, adding the results, and noting that (3•13) gives

$$\left[G_{\ell i}(t,x)\zeta_i(x) - F_{\ell i}(t,x)\eta_i(x)\right]_{x=x_1}^{x=x_2} = 0$$

as a consequence of $S_\rho(\eta,\zeta) = S_\rho(G, F) = 0$, we shall have

$$F_{\ell i}(t,x)\eta_i(x)\Big|_{x=t+0}^{x=t-0} = \eta_\ell(t) = \int_{x_1}^{x_2} G_{\ell i}(t,x)g_i(x)\,dx.$$

The formula (8•4) then follows from this by a change of notation.

Remark. The above lemma shows the connection between our boundary value system and the theory of integral equations. Indeed by defining n further sets of functions $K_{\ell i}(t, x)$, $H_{\ell i}(t, x)$ satisfying the conditions

$$\begin{aligned} &K_{\ell i}(t,t-0) - K_{\ell i}(t,t+0) = \delta_{\ell i}, \\ &\frac{\partial}{\partial x}K_{\ell i}(t,x) = L_{ik}K_{\ell k}(t,x) + M_{ik}H_{\ell k}(t,x), \\ &\frac{\partial}{\partial x}H_{\ell i}(t,x) = N_{ik}K_{\ell k}(t,x) - L_{ki}H_{\ell k}(t,x), \\ &S_\rho[K(t,x), H(t,x)] = 0, \end{aligned} \tag{8•6}$$

and having no discontinuities other than those expressly exhibited, we can prove that, <u>corresponding to functions</u> $f_i(x)$, $g_i(x)$ <u>of class</u> D, <u>every continuous solution</u> η_i, ζ_i <u>of the system</u>

$$\begin{aligned} \eta_i' &= L_{ik}\eta_k + M_{ik}\zeta_k - f_i, \\ \zeta_i' &= N_{ik}\eta_k - L_{ki}\zeta_k - g_i, \end{aligned} \quad S_\rho(\eta,\zeta) = 0, \tag{8•7}$$

is given by

$$\begin{aligned}\eta_i(x) &= \int_{x_1}^{x_2}\left[-F_{i\ell}(x,t)f_\ell(t) + G_{i\ell}(x,t)g_\ell(t)\right]dt,\\ \zeta_i(x) &= \int_{x_1}^{x_2}\left[\;H_{i\ell}(x,t)f_\ell(t) - K_{i\ell}(x,t)g_\ell(t)\right]dt,\end{aligned} \tag{8·8}$$

and the converse is also true. The coefficients of $f_\gamma(t)$, $g_\gamma(t)$ in (8·8) give the complete Green's matrix for the system (8·7) and have the properties

$$F_{i\ell}(x,t) + K_{\ell i}(t,x) = 0,\quad G_{i\ell}(x,t) = G_{\ell i}(t,x),\quad H_{i\ell}(x,t) = H_{\ell i}(t,x), \tag{8·9}$$

which are easily proved by integrating relations analogous to (8·5). The direct part of the above theorem is then proved in the same way as we proved Lemma 8·1. The converse part is proved by differentiation under sign of integration and using the discontinuity properties and the relations (8·9). Since we have no need of this theorem, the details of the proof will be omitted. We simply note that the properties (8·9) of the Green's matrix are similar to those of the matrix formed by the coefficients in (8·7).

Theorem 8·1. *For every continuous solution* η_i, ζ_i *of the system* (6·6) *corresponding to functions* g_i *of class* D, *the series*

$$u_i(x) = \sum_{\nu=1}^{\infty} \frac{c_\nu}{\sigma_\nu}\eta_{i\nu}(x),\quad \text{where}\quad c_\nu = \int_{x_1}^{x_2}\eta_{k\nu}(t)g_k(t)\,dt, \tag{8·10}$$

converge absolutely-uniformly on x_1x_2 *and represent the functions* $\eta_i(x)$ *on that interval. The functions* $\eta_{i\nu}(x)$ *are the ortho-normalized characteristic functions of* (5·1) *as described in Theorem* 7·4 *and* σ_ν *are the corresponding characteristic numbers.*

By applying (6·9) to the systems (6·6) and (5·1) we see that the series (8·10) can be written as

(8.11) $$u_i(x) = \sum_{\nu=1}^{\infty} d_\nu \eta_{i\nu}(x)\ , \quad \text{where} \quad d_\nu = \int_{x_1}^{x_2} \eta_{k\nu}(t)\eta_k(t)\,dt.$$

By applying (8.4) to the functions $\eta_{i\nu}(x)$, the series (8.10) can also be written as

(8.12) $$u_i(x) = \sum_{\nu=1}^{\infty} c_\nu \vartheta_{i\nu}(x), \quad \text{where} \quad \vartheta_{i\nu}(x) = \int_{x_1}^{x_2} G_{i\ell}(x,t)\eta_{\ell}(t)\,dt.$$

With this form of the series we can proceed to show their absolutely-uniform convergence. First for any positive integer r the inequality

$$\int_{x_1}^{x_2}\Big[g_i(t) - \sum_{\nu=1}^{r} c_\nu \eta_{i\nu}(t)\Big]\Big[g_i(t) - \sum_{\nu=1}^{r} c_\nu \eta_{i\nu}(t)\Big]dt \geq 0$$

implies the Bessel inequality

$$\sum_{\nu=1}^{r} c_\nu^2 \leq \int_{x_1}^{x_2} g_i(t) g_i(t)\,dt ,$$

which shows the convergence of $\sum_{\nu=1}^{\infty} c_\nu^2$. Next let

$$\theta_\nu(x) = \int_{x_1}^{x_2} h_\ell(x,t)\eta_{\ell\nu}(t)\,dt,$$

where $h_\ell(x,t)$ are some functions bounded in (x,t) and integrable in t for each value of x between x_1 and x_2. Given any fixed value of x and two positive integers r and s, consider the sum

$$\sum_{\nu=r}^{r+s} \big| c_\nu \theta_\nu(x)\big| = \Sigma\, c_\rho\, \theta_\rho(x) - \Sigma\, c_\tau \theta_\tau(x),$$

where the index ρ ranges over those numbers for which $c_\rho\theta_\rho(x)$ is positive and τ over those for which $c_\tau\theta_\tau(x)$ is negative. By the well-known Schwarz inequality we readily have

$$\begin{aligned}\Sigma\, c_\rho \theta_\rho(x) &= \int_{x_1}^{x_2} h_\ell(x,t)\Big[\Sigma\, c_\rho \eta_{\ell\rho}(t)\Big]dt \\ &\leq \Big[\int_{x_1}^{x_2} h_\ell(x,t) h_\ell(x,t)\,dt\Big]^{1/2}\Big[\Sigma\, c_\rho^2\Big]^{1/2}\end{aligned}$$

Since $h_\ell(x,t)$ is bounded in (x,t) and $\sum_{\nu=1}^{\infty} c_\nu$ is convergent, this inequality shows that for a sufficiently large value of r

the sum $\sum c_\rho \theta_\rho(x)$ can be made less than any assigned quantity for all values of s and x. A similar conclusion holds for the other sum $-\sum c_\tau \theta_\tau(x)$. The series $\sum_{\nu=1}^{\infty} c_\nu \theta_\nu(x)$ is therefore absolutely-uniformly convergent on x_1x_2. As the functions $G_{i\chi}(x,t)$, for each value of i, have the properties of $h\chi(x,t)$, it follows that each of the series (8·12) converges absolutely-uniformly on x_1x_2.

To show that $u_i(x)$ equals $\eta_i(x)$ we turn to the form (8·11) of the series. By the uniform convergence of these series we see that the differences $\eta_i - u_i$ are all continuous functions of x, and by termwise integration we see that they satisfy the condition

$$\int_{x_1}^{x_2} (\eta_i - u_i)\eta_{i\nu}\,dx = 0$$

for every value of ν, and hence the same condition with $\eta_{i\nu}$ replaced by u_i. By Theorem 7·3 they also satisfy the same condition with $\eta_{i\nu}$ replaced by η_i. Consequently we have

$$\int_{x_1}^{x_2} (\eta_i - u_i)(\eta_i - u_i)\,dx = 0 .$$

This shows that $\eta_i(x) - u_i(x)$ are all identically zero on x_1x_2 as was to be shown.

CHAPTER III.

SUFFICIENT CONDITIONS FOR A MINIMUM.

9. __Two auxiliary lemmas.__ In the next section we wish to make an application of the expansion theorem obtained in § 8 to the solution of Problem 2. That theorem requires the functions η_i, ζ_i to be a continuous solution of the system (6•6) corresponding to functions g_1 of class D. Consequently the functions ζ_i must be of class D' and the boundary conditions $S_\rho(\eta,\zeta) = 0$ must be satisfied. But an admissible arc η_i for Problem 2 is only a solution of the equations

$$\eta_i' = L_{ik}\eta_k + M_{ik}\zeta_k \tag{9•1}$$

corresponding to functions ζ_i which may have discontin ities and are not necessarily differentiable. The boundary conditions $S_\rho(\eta,\zeta) = 0$ also are not necessarily satisfied, if we only know that η_i satisfy the end conditions $\Psi_\mu(\eta) = 0$. In order to meet these difficulties, we have to appeal to the following Lemma 9•2 of which Lemma 9•1 is a preliminary.

__Lemma__ 9•1. __Let__ x_3x_4 __be any subinterval of__ x_1x_2. __Let__ ζ_{i0} __be any functions of__ x __continuous on__ x_3x_4, __and__ η_{i0} __the corresponding solution of__ (9•1) __with initial values__ $\eta_{i0}(x_3) = 0$. __Moreover let__ $\zeta_{i\ell} = (x - x_3)(x - x_4)\, z_{i\gamma}(x)$, $(\gamma = 1, \cdots, n)$, __where__ $z_{i\gamma}$ __denote any functions of class__ C' __on__ x_3x_4, __and__ $\eta_{i\ell}$ __be the corresponding solutions of__ (9•1) __with initial values__ $\eta_{i\ell}(x_3) = 0$. __We can always determine the functions__ $z_{i\gamma}$ __so that the matrices__ $\| \eta_{i\ell}(x_4) \|$ __and__ $\| \eta_{i0}(x_4), \eta_{i\ell}(x_4) \|$ __have the same rank for all choices of__ ζ_{i0}.

For first selecting any n fundamental solutions of the equations (9·1) with ζ_i all identically zero and then solving (9·1) in its general form by the method of variation of parameters, we see that $\eta_{i0}(x_4)$ are given by equations of the form

$$(9\cdot2) \qquad \eta_{i0}(x_4) = \int_{x_3}^{x_4} f_{ik}(x)\zeta_{k0}(x)\,dx,$$

where f_{ik} are some functions independent of ζ_{k0}. Likewise, if $\zeta_i = (x - x_3)(x - x_4)z_i(x)$, where z_i denote any functions of class C' on x_3x_4, and η_i be the corresponding solution of (9·1) with initial values $\eta_i(x_3) = 0$, we must have

$$(9\cdot3) \qquad \eta_i(x_4) = \int_{x_3}^{x_4} (x-x_3)(x-x_4) f_{ik}(x)\, z_k(x)\,dx.$$

Let $\eta_{\alpha 0}(x_4)$, where α ranges over certain of the numbers 1 to n, be a maximum subset of $\eta_{i0}(x_4)$ such that $\eta_{\alpha 0}(x_4)$ satisfy no linear relation with constant coefficients independent of ζ_{k0}. I say the numbers $\eta_\alpha(x_4)$, like $\eta_{\alpha 0}(x_4)$, also cannot satisfy a linear relation with constant coefficients independent of z_k. For if there were such a relation $c_\alpha \eta_\alpha(x_4) = 0$ with c_α not all zero, we would have by (9·3)

$$(9\cdot4) \qquad \int_{x_3}^{x_4} (x-x_3)(x-x_4)\, c_\alpha f_{\alpha k}(x)\, z_k(x)\,dx = 0$$

for all functions $z_k(x)$ of class C' on x_3x_4. If the coefficients in (9·1) are of class C', the functions $c_\alpha f_{\alpha k}(x)$ are also of class C'. By taking $z_k(x)$ equal to these functions, the relation (9·4) would require all these functions to be identically zero on x_3x_4. If the coefficients in (9·1) are only continuous so that the functions $c_\alpha f_{\alpha k}(x)$ are continuous but not necessarily of class C', we can take $z_k(x) = g_k(x, \epsilon)$ to be any functions of class C', depending upon a small number ϵ, and approaching

$c_\alpha f_{\alpha k}(x)$ uniformly on x_3x_4 as ϵ approaches zero. The same relation (9·4) would still require all the functions $c_\alpha f_{\alpha k}(x)$ to be identically zero on x_3x_4. By (9·2) the relation $c_\alpha \eta_{\alpha o}(x_4) = 0$ would then be true independent of $\zeta_{k o}$, which is a contradiction. The truth of our lemma then follows immediately.

<u>Lemma</u> 9·2. <u>Let</u> π_{i1}, π_{i2} <u>be any given constants. Let</u> $\eta_i(x)$, $\zeta_i(x)$ <u>be any solution of</u> (9·1) <u>with</u> $\eta_i(x)$ <u>continuous on</u> x_1x_2 <u>and</u> $\zeta_i(x)$ <u>of class</u> D. <u>Corresponding to any given and sufficiently small positive number</u> ϵ, <u>let each value</u> x_α, <u>for which some functions</u> $\zeta_i(x)$ <u>are discontinuous, be covered by an interval</u> $I_{\alpha\epsilon}$ <u>of length</u> 2ϵ <u>and with</u> x_α <u>as its middle point. Then there exist a positive constant</u> K <u>independent of</u> ϵ <u>and a solution</u> $\bar{\eta}_i(x, \epsilon)$, $\bar{\zeta}_i(x, \epsilon)$ <u>of</u> (9·1) <u>depending upon</u> ϵ <u>such that</u>

1) $\bar{\eta}_i(x, \epsilon)$ <u>are of class</u> C' <u>on</u> x_1x_2 <u>and</u> $\bar{\zeta}_i(x, \epsilon)$ <u>are of class</u> D';

2) $\bar{\eta}_i(x_1, \epsilon) = \eta_i(x_1)$, $\bar{\eta}_i(x_2, \epsilon) = \eta_i(x_2)$,
$\bar{\zeta}_i(x_1, \epsilon) = \pi_{i1}$, $\bar{\zeta}_i(x_2, \epsilon) = \pi_{i2}$;

3) $|\bar{\eta}_i(x, \epsilon) - \eta_i(x)| < K\epsilon$ <u>on</u> x_1x_2;
$|\bar{\zeta}_i(x, \epsilon) - \zeta_i(x)| < K$ <u>on</u> x_1x_2 <u>and</u> $< K\epsilon$ <u>on</u> x_1x_2 <u>with the small intervals</u> $I_{\alpha\epsilon}$ <u>deleted</u>.

First consider any subinterval x_3x_4 of x_1x_2 for which all the functions $\zeta_i(x)$ are continuous. Let π_{i3}, π_{i4} be any given constants. By the theory of the approximation of functions, we can determine a positive constant G independent of ϵ, and functions $\zeta_{io}(x, \epsilon)$ depending upon ϵ, satisfying the inequalities

(9·5) $$|\zeta_{io}(x,\epsilon)| < G \text{ on } (x_3x_4) \text{ and } < \epsilon \text{ on } (x_3+\epsilon, x_4-\epsilon),$$

and making the resulting functions $\zeta_i(x) + \zeta_{io}(x, \epsilon)$ of class C' on x_3x_4 and having π_{i3}, π_{i4} as end values. Let $\eta_{io}(x, \epsilon)$ denote the solution of (9.1) corresponding to $\zeta_{io}(x, \epsilon)$ and with initial values $\eta_{io}(x_3, \epsilon) = 0$. Let $\eta_{i\ell}(x)$, $\zeta_{i\ell}(x)$ be the n sets of solutions of (9.1) as determined in the preceding lemma. By that lemma we can, for each value of ϵ solve the equations

$$\eta_{io}(x_4, \epsilon) + \epsilon d_\ell(\epsilon)\eta_{i\ell}(x_4) = 0 \tag{9.6}$$

for the constants $d_\ell(\epsilon)$. Moreover by prescribing, if necessary, fixed values for some of the constants $d_\ell(\epsilon)$, it can be brought about that for each value of ϵ there is only one solution. It then follows that for each value of ϵ the functions

$$\begin{aligned} \bar{\eta}_i(x,\epsilon) &= \eta_i(x) + \eta_{io}(x,\epsilon) + \epsilon d_\ell(\epsilon)\eta_{i\ell}(x), \\ \bar{\zeta}_i(x,\epsilon) &= \zeta_i(x) + \zeta_{io}(x,\epsilon) + \epsilon d_\ell(\epsilon)\zeta_{i\ell}(x) \end{aligned} \tag{9.7}$$

are all of class C' on x_3x_4 and form a solution of (9.1) on the same interval and satisfy the end conditions

$$\bar{\eta}_i(x_3,\epsilon) = \eta_i(x_3), \quad \bar{\eta}_i(x_4,\epsilon) = \eta_i(x_4), \quad \bar{\zeta}_i(x_3,\epsilon) = \pi_{i3}, \quad \bar{\zeta}_i(x_4,\epsilon) = \pi_{i4}.$$

Next by the same reasoning as in deriving (9.2) we have

$$\eta_{io}(x,\epsilon) = \int_{x_3}^{x} f_{ik}(x,t)\zeta_{ko}(t,\epsilon)\,dt, \qquad (x_3 \le x \le x_4)$$

where f_{ik} are some functions independent of ζ_{ko} and are bounded in (x,t). This and (9.5) show that the functions $\eta_{io}(x, \epsilon)$ all approach zero with ϵ uniformly on x_3x_4. Hence (9.6) shows that the constants $d_\ell(\epsilon)$ have finite bounds as ϵ tends to zero. By (9.5), (9.7) and these results it follows readily that there exists a positive constant K independent of ϵ such that

$$|\bar{\eta}_i(x,\epsilon)-\eta_i(x)| < K\epsilon \text{ on } (x_3x_4)$$
$$|\bar{\zeta}_i(x,\epsilon)-\zeta_i(x)| < K \text{ on } (x_3x_4) \text{ and } < K\epsilon \text{ on } (x_3+\epsilon,\, x_4-\epsilon).$$

The lemma then follows immediately by applying the above method to each subinterval for which all the functions $\zeta_i(x)$ are continuous and choosing the arbitrary values like π_{i3}, π_{i4}, ... so that $\bar{\zeta}_i(x, \epsilon)$ are continuous on x_1x_2.

10. A complete solution of Problem 2. By means of the expansion theorem and Lemma 9•2 we can now prove the following

Theorem 10•1. If there is no smallest one among the characteristic numbers σ_ν of the system (5•1), the expression $J(\eta)$ in Problem 2 has no minimum. If σ_0 is the smallest one of these numbers, then the minimum of $J(\eta)$ is equal to σ_0 and a set of linearly independent minimizing arcs is given by the orthonormalized characteristic functions $\eta_i(x)$ of (5•1) corresponding to σ_0 as described in Theorem 7•4.

It is sufficient to show that the minimum of $J(\eta)$ is equal to σ_0, if it exists. The rest of the theorem is justified by the formula (5•10). For convenience we shall let σ_0 equal σ_1 of the numbers σ_ν and consider first the case $\sigma_\nu \neq 0$ for every ν. Let η_i be an admissible arc for Problem 2 and satisfy the end and norming conditions

$$(10\cdot1) \qquad \Psi_\mu(\eta)=0, \quad N(\eta)=\int_{x_1}^{x_2}\eta_i\eta_i\,dx=1.$$

We know that η_i together with some functions ζ_i of class D on x_1x_2 form a solution of (9•1). By referring to (3•8), (3•9) and noting that $S_\mu(\eta,\zeta) = 0$ is another way of writing $\Psi_\mu(\eta) = 0$, it follows that we can find solutions $\bar{\zeta}_{i1}$, $\bar{\zeta}_{i2}$ of the equations $S_\rho(\eta,\bar{\zeta}) = 0$. Let π_{i1}, π_{i2} be one of the solutions,

ϵ be any given small positive number, and $\bar{\eta}_i(x, \epsilon)$, $\bar{\zeta}_i(x, \epsilon)$ be constructed as in Lemma 9•2 with π_{i1}, π_{i2} as the end values of $\bar{\zeta}_i(x, \epsilon)$. By differentiating $\bar{\zeta}_i(x, \epsilon)$ with respect to x, we see that $\bar{\eta}_i(x, \epsilon)$, $\bar{\zeta}_i(x, \epsilon)$ form a continuous solution of the system (6•6) corresponding to some functions $\bar{g}_i(x, \epsilon)$ of class D. We can therefore apply the formulas (8•10), (8•11) of the expansion theorem and obtain

$$(10{\cdot}2) \qquad \begin{aligned} &\bar{\eta}_i = \sum_{\nu=1}^{\infty} (c_\nu(\epsilon)/\sigma_\nu)\eta_{i\nu}, \text{ where} \\ &c_\nu(\epsilon) = \int_{x_1}^{x_2} \bar{g}_i \eta_{i\nu} \, dx = \sigma_\nu \int_{x_1}^{x_2} \bar{\eta}_i \eta_{i\nu} \, dx \end{aligned}$$

Since these series all converge absolutely-uniformly on x_1x_2, we have by multiplication and termwise integration

$$(10{\cdot}3) \qquad \int_{x_1}^{x_2} \bar{\eta}_i \bar{\eta}_i \, dx = \sum_{\mu,\nu=1}^{\infty} \frac{c_\mu(\epsilon)c_\nu(\epsilon)}{\sigma_\mu \sigma_\nu} \int_{x_1}^{x_2} \eta_{i\mu}\eta_{i\nu} \, dx = \sum_{\nu=1}^{\infty} \frac{c_\nu^2(\epsilon)}{\sigma_\nu^2}.$$

Since $\bar{\eta}_i(x, \epsilon)$ approaches $\eta_i(x)$ uniformly on x_1x_2 as ϵ approaches zero and since $\eta_i(x)$ satisfy the norming condition in (10•1), the relation (10•3) on passing to limit gives

$$(10{\cdot}4) \qquad 1 = \int_{x_1}^{x_2} \eta_i \eta_i \, dx = \underset{\epsilon=0}{L} \sum_{\nu=1}^{\infty} \frac{c_\nu^2(\epsilon)}{\sigma_\nu^2}.$$

From the inequalities given in Lemma 9•2 and the expressions (5•7) to (5•9), it is also evident that

$$J(\eta) = \underset{\epsilon=0}{L} J(\bar{\eta}) = \underset{\epsilon=0}{L} \int_{x_1}^{x_2} \bar{g}_i \bar{\eta}_i \, dx.$$

By substituting $\bar{\eta}_i$ from (10•2) and termwise integration, we have

$$(10{\cdot}5) \qquad J(\eta) = \underset{\epsilon=0}{L} \int_{x_1}^{x_2} \bar{g}_i \left(\sum_{\nu=1}^{\infty} \frac{c_\nu(\epsilon)}{\sigma_\nu}\eta_{i\nu}\right) dx = \underset{\epsilon=0}{L} \sum_{\nu=1}^{\infty} \frac{c_\nu^2(\epsilon)}{\sigma_\nu}.$$

This relation and (10•4) give

$$J(\eta) - \sigma_1 = \lim_{\epsilon=0} \sum_{\nu=1}^{\infty} \frac{c_\nu^2(\epsilon)}{\sigma_\nu^2}(\sigma_\nu - \sigma_1) \geq 0 \tag{10·6}$$

as was to be proved.

If $\sigma_\nu = 0$ for some ν, let c be any small constant and consider the problem of minimizing

$$\bar{J}(\eta) = 2q(\eta) + \int_{x_1}^{x_2} [(P_{ik} + c\delta_{ik})\eta_i\eta_k + 2Q_{ik}\eta_i\eta_k' + R_{ik}\eta_i'\eta_k'] dx$$

with respect to the same functions as before. By referring to the transformations in § 3, we see that the boundary value problem for this new problem consists of the system (5·1) with N_{ik} replaced by $N_{ik} + c\,\delta_{ik}$. The new characteristic numbers are evidently the numbers $\sigma_\nu + c$. By suitably choosing c it can be brought about that no one of them is zero. The relation (10·6) then becomes $J(\eta) + c - (\sigma_1 + c) \geq 0$ which completes the proof.

The following corollary is immediate.

Corollary. <u>A necessary and sufficient condition for the second variation $I_2(\xi, \eta)$ for E_{12} to be always positive (or non-negative) with respect to all admissible variations ξ_1, ξ_2, $\eta_i(x)$, not all identically zero and satisfying the conditions $\Psi_\mu(\xi, \eta) = 0$, is that the characteristic numbers of its accessory boundary value problem be all positive (or non-negative).</u>

We shall now derive a criterion for the system (5·1) to have a least characteristic number. Problem 2 will be said to satisfy the Clebsch necessary condition, if at every point of x_1x_2 we have $R_{ik}\pi_i\pi_k \geq 0$ for all constants $(\pi) \neq (0)$ and such that $\varphi_{\alpha i}\pi_i = 0$. Under this condition we shall have

$$M_{ik}\zeta_i\zeta_k \geq 0 \tag{10·7}$$

<u>at every point of</u> x_1x_2 <u>and for all functions</u> ζ_i. For, since the determinant R in (2·9) is different from zero on x_1x_2, we can solve the equations

$$R_{ik}\pi_k+\bar{\varphi}_{i\beta}\lambda_\beta=\zeta_i\ ,\quad \varphi_{\alpha k}\pi_k=0$$

for the functions π_k and λ_β. From the definition of M_{ik} in (3·3), it is then easily verified that $M_{ik}\zeta_i\zeta_k = R_{ik}\pi_i\pi_k$ from which (10·7) follows immediately.

<u>Theorem</u> 10·2. <u>The Clebsch necessary condition for Problem</u> 2 <u>is both necessary and sufficient for the system</u> (5·1) <u>to have a least characteristic number</u>.

The necessity of the condition follows at once from Theorem 10·1 and the fact that it is necessary for Problem 2 to have a minimum. To prove its sufficiency, we proceed as follows: In the theory of quadratic forms, it is well known that any quadratic form $a_{ik}x_ix_k$ will be $\geq cx_ix_i$, if c is a sufficiently large and negative number. For the signs of the chain of principal minors of the determinant $a_{ik} - c\,\delta_{ik}$, beginning from the upper left hand corner, depend upon those of $-c$, c^2, $-c^3$, ···, if the absolute value of c is sufficiently large. Hence, for c sufficiently large and negative, these signs are all positive and the form $(a_{ik} - c\,\delta_{ik})x_ix_k$ is positive definite. From this theorem it follows that there exists a constant c such that $2q(\eta)$ is always $\geq c(\eta_{i1}\eta_{i1} + \eta_{i2}\eta_{i2})$ and hence $J(\eta)$ is always greater than or equal to

$$\bar{J}(\eta)=c(\eta_{i1}\eta_{i1}+\eta_{i2}\eta_{i2})+\int_{x_1}^{x_2}2\omega(x,\eta,\eta')\,dx\ .$$

Let h(x) be any suitable function which takes the value -1 at x_1 and $+1$ at x_2. The expression $\bar{J}(\eta)$ can also be written as

$$\bar{J}(\eta)=\int_{x_1}^{x_2}\left[c\frac{d}{dx}(\eta_i\eta_i h)+2\omega(x,\eta,\eta')\right]dx=\int_{x_1}^{x_2}(\bar{P}_{ik}\eta_i\eta_k+2\bar{Q}_{ik}\eta_i\eta_k'+R_{ik}\eta_i'\eta_k')dx,$$

where $\bar{P}_{ik}$, $\bar{Q}_{ik}$ are some functions having the properties of P_{ik}, Q_{ik}. By the transformations by which we arrived at (5·7) we see that, for an admissible arc η_i, ζ_i of Problem 2, the value of $\bar{J}(\eta)$ is given by

$$\bar{J}(\eta)=\int_{x_1}^{x_2}(\bar{N}_{ik}\eta_i\eta_k+M_{ik}\zeta_i\zeta_k)\,dx,$$

where $\bar{N}_{ik}$ are some functions having the same properties as N_{ik}. Letting l be any constant and η_i be normed, we shall have

$$\bar{J}(\eta)-l=\int_{x_1}^{x_2}\left[(N_{ik}-l\delta_{ik})\eta_i\eta_k+M_{ik}\zeta_i\zeta_k\right]dx.$$

This expression together with the above theorem on quadratic forms and the inequality (10·7) shows that $\bar{J}(\eta)$ will be always $\geq l$ provided the latter is taken sufficiently large and negative. Hence, a fortiori, $J(\eta)$ will have a lower bound and our theorem follows immediately from the formula (5·10).

From the above two theorems we can easily deduce the following theorem which gives a meaning to each of the characteristic numbers of (5·1) and is important for the development in Chapter IV.

Theorem 10·3. Let the Clebsch necessary condition be satisfied. Let the totality of characteristic numbers of (5·1) be arranged in the order $\sigma_1 \leq \sigma_2 \leq \sigma_3 \leq \cdots\cdots$, (each number being counted with a multiplicity equal to its index), and $\eta_{i\nu}$, $\zeta_{i\nu}$ be the corresponding orthonormalized characteristic solutions of (5·1) as described in Theorem 7·4. Let (H_1) denote the totality of arcs η_i, admissible for Problem 2 and satisfying the conditions (10·1). Let (H_γ) for $\gamma > 1$ denote the subclass

<u>of</u> (H_1) <u>which satisfies the additional conditions</u>

$$(10\cdot 8) \qquad \int_{x_1}^{x_2} \eta_i \eta_{is}\, dx = 0 , \qquad (s = 1, \cdots, \ell - 1).$$

<u>The minimum of the expression</u> $J(\eta)$ <u>in Problem</u> 2 <u>with respect to all arcs of the class</u> (H_χ), (χ = 1, 2, 3, ···) <u>is then equal to</u> σ_ℓ , <u>and a set of linearly independent minimizing arcs is given by the characteristic functions</u> $\eta_{i\nu}$ <u>which are in the class</u> (H_χ) <u>and with characteristic numbers</u> σ_ν <u>equal to</u> σ_ℓ .

Its proof is immediate. For let η_i be an arc of the class (H_χ). By (10·8) and (10·2) we have $\underset{\epsilon=0}{L}\, c_s(\epsilon) = 0$ for $s = 1, \cdots, \chi-1$. The relations (10·4) and (10·5) then give the desired result

$$J(\eta) - \sigma_\ell = \underset{\epsilon=0}{L} \sum_{\nu=\ell}^{\infty} \frac{c_\nu^2(\epsilon)}{\sigma_\nu^2}(\sigma_\nu - \sigma_\ell) \geq 0 .$$

11. <u>Sufficient conditions for Problem</u> 1. In this section we shall deduce a set of sufficient conditions for Problem 1 as formulated in § 1. In that section we have let E_{12} denote the particular admissible arc whose minimizing property is to be studied. For the purpose of distinguishing and convenience, we shall now let $\overline{E}$ denote this particular arc and $(\overline{x}_1, \overline{y}_1, \overline{x}_2, \overline{y}_2) = (\overline{x}_1, \overline{y}_{i1}, \overline{x}_2, \overline{y}_{i2})$ represent its end points. The notation E will denote any extremal arc with end points (x_1, y_1, x_2, y_2) in some neighborhood $(\overline{x}_1, \overline{y}_1, \overline{x}_2, \overline{y}_2)_\delta$ and I(E) the value of I given in (1·4) evaluated for E. Simil[arly] I(C) will denote the value of I evaluated for any admissible arc C. The arc $\overline{E}$ will be said to satisfy the <u>Weierstrass sufficient condition</u> if the function

$$\mathcal{E}(x, y, y', Y', \lambda) = F(x, y, Y', \lambda) - F(x, y, y', \lambda) - (Y_i' - y_i')F_{y_i'}(x, y, y', \lambda),$$

where F is defined by (1•10) with $\lambda_0 = 1$, is positive for every set of elements (x, y, y', Y', λ) for which the set (x, y, y', λ) is in a neighborhood of similar sets belonging to $\bar{E}$ and $(x, y, Y') \neq (x, y, y')$ is admissible. The sufficiency theorem is then as follows:

Theorem 11•1. For a non-singular extremal arc $\bar{E}$ of Problem 1 to afford a proper strong relative minimum, it is sufficient that 1) its end points satisfy the conditions $\psi_\mu = 0$ but no other pair of points on it does so, 2) its end points satisfy the transversality conditions (1•12), 3) it is absolutely normal on every subinterval $x_1 x_3$ (according to the definition on page 6), 4) it satisfies the Weierstrass sufficient condition, and 5) the characteristic numbers of its accessory boundary value problem are all positive.

In order to prove the theorem we first give the following preliminary facts. Let (a, b) stand for the set (a_i, b_i) and $z_i = F_{y_i'}$. Since $\bar{E}$ is a non-singular extremal, it is well known that there exists[1] a 2n-parameter family of extremals

$$\begin{aligned} &y_i = y_i(x, a, b), \\ (11\cdot1) \qquad &z_i = z_i(x, a, b), \quad (a, b) \text{ in some neighborhood } (\bar{a}, \bar{b})_\epsilon, \\ &\lambda_\alpha = \lambda_\alpha(x, a, b), \end{aligned}$$

which contains $\bar{E}$ for $(a, b) = (\bar{a}, \bar{b})$. The functions y_i, y_{ix}, z_i, z_{ix}, λ_α are all of class C'' in a neighborhood of the values (x, a, b) belonging to $\bar{E}$ and satisfying the conditions

$$(11\cdot2) \qquad a_i = y_i(\bar{x}_1, a, b), \quad b_i = z_i(\bar{x}_1, a, b).$$

[1] Bliss, The problem of Lagrange, loc.cit., pp. 685 - 687.

Having had these facts we then base our proof of the theorem upon the following three lemmas which hold under the hypotheses of the theorem.

Lemma 11·1. There exist two positive numbers ϵ and $\delta\,(\leq \epsilon)$ such that if the subscript ϵ in (11·1) is taken equal to this ϵ, then through every pair of points (x_1, y_1, x_2, y_2) in the neighborhood $(\bar{x}_1, \bar{y}_1, \bar{x}_2, \bar{y}_2)_\delta$ there passes a unique extremal arc E of the family (11·1). The parameters $a_i(x_1, y_1, x_2, y_2)$ and $b_i(x_1, y_1, x_2, y_2)$, determining this extremal, are functions of class C" in their arguments and approach $\bar{a}_i$ and $\bar{b}_i$ respectively as δ tends to zero.

Its proof follows immediately from the usual implicit function theorem if we notice that the equations

$$y_{i1} = y_i(x_1, a, b), \quad y_{i2} = y_i(x_2, a, b)$$

have the initial solution $(\bar{x}_1, \bar{y}_1, \bar{x}_2, \bar{y}_2, \bar{a}, \bar{b})$ and the determinant

$$\Delta = \begin{vmatrix} y_{ia_k}(\bar{x}_1, \bar{a}, \bar{b}), & y_{ib_k}(\bar{x}_1, \bar{a}, \bar{b}) \\ y_{ia_k}(\bar{x}_2, \bar{a}, \bar{b}), & y_{ib_k}(\bar{x}_2, \bar{a}, \bar{b}) \end{vmatrix}$$

is different from zero. The justification of $\Delta \neq 0$ is also evident under the conditions 3) and 5) of the theorem. For from the fact that (11·1) gives a 2n-parameter family of extremals satisfying the conditions (11·2) and containing $\bar{E}$ for $(a, b) = (\bar{a}, \bar{b})$, it follows that the 2n sets of functions

$$y_{ia_k}(x, \bar{a}, \bar{b}), \ z_{ia_k}(x, \bar{a}, \bar{b}) \ \text{and} \ y_{ib_k}(x, \bar{a}, \bar{b}), \ z_{ib_k}(x, \bar{a}, \bar{b})$$

form 2n linearly independent solutions of the accessory equations set up for $\bar{E}$ and hence any linear combination of them

with constant coefficients not all zero is a non-identically vanishing solution of the same equations. If Δ were zero, there would be such a linear combination, say $\eta_i(x)$, $\zeta_i(x)$, with $\eta_i(x)$ all vanishing at x_1 and x_2 but not all identically zero on x_1x_2. By taking $\xi_1 = \xi_2 = 0$ with the functions $\eta_i(x)$ to form a set of admissible variations and by arguments similar to the proof of Lemma 5·2, we would have $I_2(\xi, \eta) = 0$, which is in contradiction with the corollary to Theorem 10·1.

Lemma 11·2. *If ϵ is given by Lemma 11·1 and δ is taken sufficiently small, then $I(E) > I(\bar{E})$ for every extremal arc E of the family (11·1), distinct from $\bar{E}$, with end points in $(\bar{x}_1, \bar{y}_1, \bar{x}_2, \bar{y}_2)_\delta$, and satisfying the conditions $\psi_\mu = 0$.*

For convenience we shall use the notation u_s, $(s = 1, \cdots, 2n+2)$ for the set $(x_1, y_{i1}, x_2, y_{i2})$. By Lemma 11·1 the function $I(E)$ for extremal arcs as described in the lemma may be regarded as a function $I(u)$ of the variables u_s defined in a neighborhood of $\bar{u}_s$ and subject to the conditions $\psi_\mu = 0$. Let $\lambda_0 = 1$, d_μ be the set of constants occurring in the transversality conditions (1·12), $\vartheta(u)$ denote the sum $I(u) + d_\mu \psi_\mu(u)$, and $\bar{\vartheta}_s$, $\bar{\vartheta}_{st}$, $\bar{\psi}_{\mu s}$ represent the partial derivatives ϑ_{u_s}, $\vartheta_{u_s u_t}$, $\psi_{\mu u_s}$, $(s, t = 1, \cdots, 2n+2)$ evaluated at $u_s = \bar{u}_s$. With the help of (1·11) and (1·12) it is easy to verify that $\bar{\vartheta}_s = 0$ for every value of s. Next let v_s be $2n + 2$ new variables. If we can show that $\bar{\vartheta}_{st} v_s v_t > 0$ for all sets $(v) \neq (0)$ and satisfying the conditions $\bar{\psi}_{\mu s} v_s = 0$, then, by the theory of ordinary maxima and minima, it will follow that $I(u) > I(\bar{u})$ for all sets (u) sufficiently near but distinct from $(\bar{u})$ and satisfying the conditions $\psi_\mu(u) = 0$.

In order to show that $\bar{\vartheta}_{st} v_s v_t > 0$ for all sets (v)

as mentioned above, we proceed as follows: Let v_s be any such set. Since the matrix $\|\overline{\psi}_{\mu s}\|$ is assumed to be of rank $p \leq 2n + 2$ we may, for convenience, let the determinant $|\overline{\psi}_{\mu\nu}|$, $(\mu, \nu = 1, \cdots, p)$ be different from zero. If $p = 2n + 2$, Problem 1 is a problem with fixed end points, in which case we do not need the present lemma. If $p < 2n + 2$, we let $u_r(e) = ev_r + \overline{u}_r$, $(r = p+1, \cdots, 2n+2)$ and consider the equations

$$\psi_\mu[u_1, \cdots, u_p, u_{p+1}(e), \cdots, u_{2n+2}(e)] = 0 . \tag{11·3}$$

Since they have the initial solution $(u_1, \cdots, u_p, e) = (\overline{u}_1, \cdots, \overline{u}_p, 0)$ and the determinant $|\overline{\psi}_{\mu\nu}|$ is different from zero, there exists a solution $u_s(e)$, $(s = 1, \cdots, 2n+2)$ of class C'' in a neighborhood of $e = 0$ and reducing to $\overline{u}_s$ for $e = 0$. By differentiating (11·3) with respect to e, putting $e = 0$, and noting that v_s is a given solution of $\overline{\psi}_{\mu s} v_s = 0$, it follows that $u_{se}(0) = v_s$. By Lemma 11·1 it also follows that for each sufficiently small value of e there is a unique extremal arc E

$$y_i(x,e) = y_i[x, a(e), b(e)] , \quad \lambda_\alpha(x,e) = \lambda_\alpha[x, a(e), b(e)] \tag{11·4}$$

of the family (11·1) with $u_s(e)$ as end values. As e varies, these arcs E form a one-parameter family of extremals, satisfying the conditions $\psi_\mu = 0$ and containing $\overline{E}$ for $e = 0$. The function $I[u(e)] = I(E)$ is therefore identical with the function

$$\mathcal{J}[u(e)] = I[u(e)] + d_\mu \psi_\mu[u(e)] .$$

From this and the fact that $\overline{\mathcal{J}}_s = 0$, $u_{se}(0) = v_s$, we have

$$\frac{d^2}{de^2} I[u(e)]\Big|^{e=0} = \frac{d^2}{de^2}\mathcal{J}[u(e)]\Big|^{e=0} = \overline{\mathcal{J}}_{st} v_s v_t \tag{11·5}$$

On the other hand, by referring to (2·2), we see that the

expression (11·υ) is also equal to $I_2(\xi, \eta)$, where ξ_1, ξ_2, $\eta_i(x)$ are the variations of the family (11·4). Since $(v) \neq (0)$, the set ξ_1, ξ_2, $\eta_i(x)$ cannot be all identically zero. Therefore $\mathcal{J}_{st} v_s v_t > 0$ by the condition 5) of the theorem and the corollary to Theorem 10·1.

Lemma 11·3. If ϵ is given by Lemma 11·1 and δ is taken sufficiently small, there will exist a neighborhood F of $\overline{E}$ such that every extremal arc E of the family (11·1) with end values in $(\overline{x}_1, \overline{y}_1, \overline{x}_2, \overline{y}_2)_\delta$ lies in F and has the property that I(C) > I(E) for every admissible arc C in F joining the ends of E but distinct from E.

With the conditions 3) and 5) in the theorem we can show[1] that there exists a conjugate system of solutions $\eta_{ik}(x)$, $\zeta_{ik}(x)$ of the accessory equations set up for $\overline{E}$ such that the determinant $|\eta_{ik}(x)|$ is everywhere different from zero on $\overline{x}_1\overline{x}_2$. Let a_i, b_i, c_i be any 3n parameters and

$$(11·6) \qquad y_i = Y_i(x,a,b,c), \quad z_i = Z_i(x,a,b,c), \quad \lambda_\alpha = \Lambda_\alpha(x,a,b,c)$$

be the results of replacing a_i by $a_i + c_k \eta_{ik}(\overline{x}_1)$ and b_i by $b_i + c_k \zeta_{ik}(\overline{x}_1)$ in equations (11·1). As a consequence of (11·2), the functions Y_i, Z_i will satisfy the conditions

$$(11·7) \qquad Y_i(\overline{x}_1, a, b, c) = a_i + c_k \eta_{ik}(\overline{x}_1), \quad Z_i(\overline{x}_1, a, b, c) = b_i + c_k \zeta_{ik}(\overline{x}_1).$$

By means of these conditions and the fact that (11·6) with fixed $(a, b) = (\overline{a}, \overline{b})$ and variable (c) defines an n-parameter family of extremals containing $\overline{E}$ for $(c) = (0)$, we see that the derivatives $Y_{ic_k}(x, \overline{a}, \overline{b}, 0)$, $Z_{ic_k}(x, \overline{a}, \overline{b}, 0)$ are exactly the

---

1 See Theorem 14·2 below.

functions $\eta_{ik}(x)$, $\zeta_{ik}(x)$. The equations $y_i = Y_i(x, a, b, c)$ therefore have the initial solution $(a, b, c) = (\bar{a}, \bar{b}, 0)$, (x, y) on $\bar{E}$, and the determinant $|Y_{ic_k}|$ is everywhere different from zero for this initial solution. By the extended implicit function theorem there exist two positive numbers τ and σ such that for every (a, b) in $(\bar{a}, \bar{b})_\tau$ and (x, y) in $(\bar{E})_\tau$ these equations have a unique solution $c_i(x, y, a, b)$, whose functions are of class C'', less than σ in absolute value, and approach zero with τ. Because of the uniqueness of this solution it is also evident that, when (x, y) varies on an extremal E situated in $(\bar{E})_\tau$ and determined by (11·1) with (a, b) in $(\bar{a}, \bar{b})_\tau$ the functions $c_i(x, y, a, b)$ have to be identically zero. On account of the condition 4) in the theorem we can take τ sufficiently small so that, when p_i and ℓ_α are defined by the equations

$$(11\cdot8) \quad p_i(x,y,a,b) = Y_i'[x, a, b, c(x,y,a,b)], \quad \ell_\alpha(x,y,a,b) = \Lambda_\alpha[x, a, b, c(x,y,a,b)],$$

the Weierstrass function $\mathcal{E}(x, y, p, y', \ell)$ will be positive for all (a, b) in $(\bar{a}, \bar{b})_\tau$, (x, y) in $(\bar{E})_\tau$ and all admissible sets $(x, y, y') \neq (x, y, p)$.

Now suppose that a particular τ effective in the above argument has been taken. By Lemma 11·1 we can take δ sufficiently small so that every extremal arc E of the family (11·1) with end values in $(\bar{x}_1, \bar{y}_1, \bar{x}_2, \bar{y}_2)_\delta$ has its parameter values in $(\bar{a}, \bar{b})_\tau$ and E itself lies in $(\bar{E})_\tau$. To fix the ideas, let an effective δ be taken and E_0 be any such extremal with a_{i0}, b_{i0} as its parameter values. It is then evident that the n-parameter family of extremals

$$(11\cdot9) \quad y_i = Y_i(x, a_0, b_0, c), \quad z_i = Z_i(x, a_0, b_0, c), \quad \lambda_\alpha = \Lambda_\alpha(x, a_0, b_0, c)$$

obtained from (11·6) by taking $(a, b) = (a_0, b_0)$ and subject to the conditions $|c_i| < \sigma$, will simply cover the region $(\overline{E})_\tau$, to which E_0 is interior, and include E_0 for $(c) = (0)$. Moreover the derivatives $Y_{ic_k}(x, a_0, b_0, 0)$, $Z_{ic_k}(x, a_0, b_0, 0)$ will form a conjugate system of solutions of the accessory equations set up for E_0, because by (11·7) they take on the values $\eta_{ik}(\overline{x}_1)$, $\zeta_{ik}(\overline{x}_1)$ at $x = \overline{x}_1$. The family (11·9) therefore forms a Mayer field surrounding E_0 with slope functions and multipliers given by (11·8) with $(a, b) = (a_0, b_0)$. Since the corresponding Weierstrass function is everywhere positive in this field, the usual fundamental sufficiency theorem[1] then gives the lemma with $(\overline{E})_\tau$ as the neighborhood F.

With the above three lemmas we can easily complete the proof of the theorem as follows: Suppose that a number δ and a neighborhood F of $\overline{E}$ effective in those lemmas have been chosen. We first show that there exists a neighborhood $\mathfrak{N}$ of $\overline{E}$ such that every pair of points lying in $\mathfrak{N}$ and satisfying the conditions $\psi_\mu = 0$ will be in the neighborhood $(\overline{x}_1, \overline{y}_1, \overline{x}_2, \overline{y}_2)_\delta$. Because if this were not true, then for every positive integer q there would exist a pair of points, situated in $(\overline{E})_{1/q}$, satisfying the conditions $\psi_\mu = 0$, but not in $(\overline{x}_1, \overline{y}_1, \overline{x}_2, \overline{y}_2)_\delta$. These pairs of points would have a pair of accumulation points, lying on $\overline{E}$, satisfying the conditions $\psi_\mu = 0$ but distinct from the pair $(\overline{x}_1, \overline{y}_1, \overline{x}_2, \overline{y}_2)$, in contradiction with the condition 1) of the theorem. The existence of $\mathfrak{N}$ is therefore proved. Next let $\mathfrak{N}$ be taken interior to F and C be any admissible arc for Problem 1, lying in $\mathfrak{N}$ and satisfying the

---

[1] Bliss, The problem of Lagrange, loc.cit., pp. 731 - 732.

conditions $\psi_\mu = 0$. Because the ends of C are in $(\bar{x}_1, \bar{y}_1, \bar{x}_2, \bar{y}_2)_\delta$, the above lemmas then show that through the ends of C there passes an extremal E of the family (11·1) lying in F and such that $I(C) \geq I(E) \geq I(\bar{E})$. The equality $I(C) = I(\bar{E})$ can hold only when C coincides with $\bar{E}$. Therefore $I(C) > I(\bar{E})$ whenever C is distinct from $\bar{E}$ as was to be proved.

CHAPTER IV.

OSCILLATION AND COMPARISON THEOREMS.

In this chapter we shall go back to the differential equations in (5•1). By considering certain special solutions of these equations in connection with their corresponding minimizing problems, we shall be able to deduce some rather general oscillation and comparison theorems which especially throw light on the theory of conjugate and focal points. The proofs are chiefly based upon Theorem 10•3 and consequently the Clebsch necessary condition stated on page 50 is assumed to be satisfied throughout this chapter.

12. Conjugate solutions and focal points. For any two solutions η_i, ζ_i and $\bar{\eta}_i$, $\bar{\zeta}_i$ of the differential equations in (5•1) corresponding to a fixed value of σ, the same multiplying and adding process as used in the proof of Lemma 6•2 will show that

$$(d/dx)(\eta_i\bar{\zeta}_i-\bar{\eta}_i\zeta_i)=0 \quad \text{and hence} \quad \eta_i\bar{\zeta}_i-\bar{\eta}_i\zeta_i=\text{constant}.$$

If this constant is zero, the two solutions are said to be mutually conjugate according to von Escherich. A conjugate system of solutions is a system in which any two solutions are mutually conjugate. The maximum number of linearly independent solutions in such a system is evidently n. Accordingly, we shall call a conjugate system consisting of n linearly independent solutions a conjugate base and the set of all their linear combinations with constant coefficients a conjugate family. Any conjugate base may obviously be replaced by a new base consisting

of n linearly independent linear combinations of the old base without changing the conjugate property or the family determined by the base. Now consider a conjugate base

$$\eta_i^l(x,\sigma),\ \zeta_i^l(x,\sigma), \qquad (l=1,\cdots,n) \tag{12·1}$$

with constant initial values $\eta_i^l(x_1,\sigma) = b_{\gamma i}$, $\zeta_i^l(x_1,\sigma) = a_{\gamma i}$. The rows of $\|a_{\gamma i}\|$ and $\|b_{\gamma i}\|$ must therefore be mutually conjugate. We shall denote the rank of $\|b_{\gamma i}\|$ by $n - p$, $(0 \leq p \leq n)$. If $p > 0$, we assume that the base (12·1) has been changed so that $b_{\mu i} = 0$, $(\mu = 1, \cdots, p;\ i = 1, \cdots, n)$. If for a pair of values σ_0 and x_3, $(x_1 < x_3 \leq x_2)$ the determinant

$$\Delta(x,\sigma) = |\eta_i^l(x,\sigma)| \tag{12·2}$$

is zero, some members of the family determined by (12·1) must be solutions of the system

$$\begin{aligned} \eta_i' &= L_{ik}\eta_k + M_{ik}\zeta_k, & S_l(\eta,\zeta) &= a_{li}\eta_{i1} - b_{li}\zeta_{i1} = 0, \\ \zeta_i' &= N_{ik}\eta_k - L_{ki}\zeta_k - \sigma\eta_i, & \eta_{i3} &= 0, \end{aligned} \tag{12·3}$$

for $\sigma = \sigma_0$. The focal points of the family are usually defined in connection with these solutions. But before giving a formal definition of such points, it is desirable to first point out the connection of the system (12·3) with a minimizing problem. In fact by the conjugate property of $\|a_{\gamma i}\|$, $\|b_{\gamma i}\|$ and the above agreement on their values, the argument in § 4 shows that there exists a quadratic form $A_{ik}\eta_{i1}\eta_{k1}$ such that (12·3) is the system arising from the problem of minimizing

$$J_{13}(\eta) = 2g(\eta) + \int_{x_1}^{x_3} 2\omega(x,\eta,\eta')\,dx\,, \text{ where } 2g(\eta) = A_{ik}\eta_{i1}\eta_{k1}, \tag{12·4}$$

with respect to all functions $\eta_i(x)$ of class D' on x_1x_3 and

satisfying the conditions

$$(12\cdot5)\qquad \begin{aligned} &\Phi_\alpha(x,\eta,\eta') = \varphi_{\alpha i}\eta_i' + \theta_{\alpha i}\eta_i = 0\,, && N_{13}(\eta) = \int_{x_1}^{x_3}\eta_i\eta_i\,dx = 1\,, \\ &\Psi_\mu(\eta) = a_{\mu i}\eta_{i1} = 0\,, && \eta_{i3} = 0\,, \end{aligned}$$

in which the conditions $\Psi_\mu(\eta) = 0$ are another way of writing $S_\mu(\eta,\zeta) = 0$ and the number of them is p, $(0 \le p \le n)$. If for a fixed value of x_3 the system (12•3) has no characteristic solution with $\eta_i(x)$ all identically zero on x_1x_3, it is said to be normal on that interval. In the following discussion we shall assume that it is normal on every subinterval x_1x_3. If $p = 0$, this normality condition is always satisfied as is seen by a consideration of the conditions $S_\rho(\eta,\zeta) = 0$. When $p > 0$, Lemma 5•1 shows that this is still true provided a hypothesis corresponding to the hypothesis (δ) in Problem 2 is always satisfied.

A focal point of x_1[1] relative to the conjugate family determined by (12•1) or to the system (12•3) or to the corresponding minimizing problem and corresponding to a number σ_0 may then be defined as a point x_3, if it exists, for which (12•3) with $\sigma = \sigma_0$ has at least one characteristic solution. The number of linearly independent characteristic solutions is defined as the index of the corresponding focal point. When the conditions $S_\rho(\eta,\zeta) = 0$ become $\eta_{i1} = 0$, the focal points are called conjugate points. The usual definition of focal points are those corresponding to $\sigma_0 = 0$.

Since a solution of (12•3) must be a linear combination of the functions in (12•1) with constant coefficients c_γ satis-

---

[1] We use this phrase for convenience. Strictly speaking the focal point thus defined is a focal point of the end-manifold at x_1 for the minimizing problem.

fying the conditions $c_\chi \eta_i^\ell(x_3, \sigma) = 0$ and the number of linearly independent solutions of (12•3) is exactly the number of linearly independent solutions in c_χ of these conditions, it follows that, for a pair of values x_3 and σ_0 for which $\Delta(x_3, \sigma_0)$ has rank n - r, $(0 < r \leq n)$, the number σ_0 is a characteristic number of (12•3) of index r and the point x_3 is a focal point of x_1 corresponding to σ_0 and is of the same index r. For the general system (5•1) the index of a characteristic number is ≥ 1 and $\leq 2n$. In the present case it cannot exceed n. An examination of the proof of Theorem 7•2 will show that the index of σ_0 is equal to its multiplicity as a root of $\Delta(x_3, \sigma_0) = 0$ as is the case with the determinant $D(\sigma)$ involved there. When the form $M_{ik} \zeta_i \zeta_k$ is definite, we can easily show that this index is also equal to the multiplicity of x_3 as a root of $\Delta(x_3, \sigma_0) = 0$. But when $\|M_{ik}\|$ is singular, this is not necessarily true as can be shown by example.

Now think of σ as a function of x_3 defined by the equation $\Delta(x_3, \sigma) = 0$. With the Clebsch necessary condition and the normality condition on the system (12•3) as mentioned above, it follows from our preceding theorems that for each value of x_3 the roots of $\Delta(x_3, \sigma) = 0$ are all real, denumerably infinite in number, and have a least one. For convenience in the following discussions we shall count a root of index r as r roots and arrange the totality of these roots and their corresponding orthonormalized characteristic solutions in the sequence

$$(12\cdot 6) \qquad \left\{ \sigma_\nu(x_3), \eta_i^\nu(x|x_3), \zeta_i^\nu(x|x_3) \right\} \text{ with } \sigma_1(x_3) \leq \sigma_2(x_3) \leq \cdots .$$

It is easy to see that a knowledge of the variations of the

numbers $\sigma_\nu(x_3)$ will lead to a knowledge of the distribution and number of focal points of x_1 corresponding to different values of σ.

13. Two fundamental oscillation theorems. We first prove the following

Lemma 13·1. The numbers $\sigma_\nu(x_3)$ in (12·6) all become positively infinite as x_3 tends to x_1.

First let the quadratic form $2q(\eta)$ in (12·4) be identically zero. Let c be any given number and $u_{i\gamma}$, $v_{i\gamma}$, $(\gamma = 1, \cdots, n)$ be n sets of solutions of the equations

$$
(13\cdot 1)\qquad \begin{aligned} u_i' &= L_{ik}u_k + M_{ik}v_k, \\ v_i' &= N_{ik}u_k - L_{ki}v_k - c\,u_i \end{aligned}
$$

with initial values $u_{i\gamma}(x_1) = \delta_{i\ell}$, $v_{i\gamma}(x_1) = 0$ so that they are mutually conjugate in pairs. Since $|u_{i\gamma}(x)| \neq 0$ at $x = x_1$, there is an interval $x_1, x_1 + h$ for which it is everywhere different from zero. Let x_3 be chosen so that $x_3 - x_1 \leq h$. I say $\sigma_1(x_3)$ will be $\geq c$. Indeed let σ, η_i, ζ_i be the set $\sigma_1(x_3)$, $\eta_i^{(1)}(x|x_3)$, $\zeta_i^{(1)}(x|x_3)$ given in (12·6) and solve the equations $\eta_i = u_{i\gamma}a_\gamma$ for the functions a_γ on x_1x_3. By (13·1) and the notations

$$
(13\cdot 2)\qquad v_{i\ell}a_\ell = v_i, \quad u_{i\ell}a_\ell' = U_i, \quad v_{i\ell}a_\ell' = V_i,
$$

one can readily verify that η_i, v_i satisfy the equations

$$
(13\cdot 3)\qquad \begin{aligned} \eta_i' &= L_{ik}\eta_k + M_{ik}v_k + U_i, \\ \zeta_i' &= N_{ik}\eta_k - L_{ki}v_k + V_i - c\,\eta_i. \end{aligned}
$$

Multiplying these equations respectively by ζ_i, $-\eta_i$, those in (12·3) respectively by $-v_i$, η_i, adding the results, and

integrating from x_1 to x_3, we shall have

$$(13\cdot 4) \qquad \sigma - c = \zeta_i \eta_i \Big|_3^1 + \int_{x_1}^{x_3} (\zeta_i U_i - \eta_i V_i)\, dx .$$

By the conjugate property of $u_{i\gamma}$, $v_{i\gamma}$ we have $u_{ik}v_{i\gamma} = u_{i\gamma}v_{ik}$ identically on x_1x_3. Hence by the definition of a_γ and (13•2)

$$\eta_i V_i = (u_{ik} a_k)(v_{il} a'_l) = (u_{il} a'_l)(v_{ik} a_k) = v_i U_i .$$

The differential equations in (12•3) and (13•3) also give $U_i = M_{ik}(\zeta_k - v_k)$. The relation (13•4) therefore becomes

$$\sigma - c = \zeta_i \eta_i \Big|_3^1 + \int_{x_1}^{x_3} M_{ik}(\zeta_i - v_i)(\zeta_k - v_k)\, dx$$

which is ≥ 0, because $\zeta_i \eta_i \Big|_3^1 = 0$ by a relation analogous to (3•12) and the assumption $2q(\eta) = 0$, and the integrand is ≥ 0 by (10•7).

If $2q(\eta) = A_{ik}\eta_{i1}\eta_{k1} \neq 0$, the reasoning in the proof of Theorem 10•2 shows that $2q(\eta)$ will be $\geq h\,\eta_{i1}\eta_{k1}$ for a sufficiently large and negative number h. The expression $J_{13}(\eta)$ in (12•4) will then be greater than or equal to

$$(13\cdot 5) \qquad \bar{J}_{13}(\eta) = h\eta_i\eta_i\Big|_3^1 + \int_{x_1}^{x_3} 2\omega(x,\eta,\eta')\,dx = \int_{x_1}^{x_3} \left[-h(\eta_i\eta_i)' + 2\omega(x,\eta,\eta')\right] dx$$

for all functions $\eta_i(x)$ satisfying the conditions $\eta_{i3} = 0$. Since the last integrand in (13•5) is another function $2\bar{\omega}(x,\eta,\eta')$ of the same form as $2\omega(x,\eta,\eta')$ and with the same coefficients R_{ik}, the above proof shows that the numbers $\bar{\sigma}_\nu(x_3)$ of the system arising from minimizing $\bar{J}_{13}(\eta)$ all become infinite as x_3 tends to x_1. The general truth of our lemma then follows immediately.

<u>Theorem</u> 13•1. <u>Each of the numbers</u> $\sigma_\nu(x_3)$ <u>in</u> (12•6) <u>varies continuously with</u> x_3 <u>and properly increases from</u> $\sigma_\nu(x_2)$

<u>to infinity as</u> x_3 <u>decreases from</u> x_2 <u>to</u> x_1.

In order to prove the continuity of the numbers $\sigma_\nu(x_3)$ we notice that $\Delta(x, \sigma)$ is continuous in (x, σ) and analytic in σ. At a particular fixed value x_3, let r and only r of the numbers $\sigma_\nu(x_3)$ be equal to σ_0. In the complex σ-plane there exists a circle with σ_0 as center and within which $\Delta(x_3, \sigma)$ is everywhere different from zero except at $\sigma = \sigma_0$. Letting ϵ be any positive number less than the radius of this circle and C_ϵ be the circle with σ_0 as center and ϵ as radius, we show that there exists a positive number δ such that, whenever x lies within both the intervals (x_1, x_2) and $(x_3-\delta, x_3+\delta)$, the equation $\Delta(x, \sigma) = 0$ has exactly r roots lying within C_ϵ and consequently within the interval $(\sigma_0-\epsilon, \sigma_0+\epsilon)$. In fact since $\Delta(x, \sigma)$ is continuous in (x, σ) for x near x_3 and σ on C_ϵ and is everywhere different from zero for $x = x_3$ and σ on C_ϵ, there exists a $\delta > 0$ such that, whenever x lies within both the intervals (x_1, x_2) and $(x_3-\delta, x_3+\delta)$, the function $\Delta(x, \sigma)$ is everywhere different from zero along C_ϵ. As is well known in the theory of functions of a complex variable, for x in the region just mentioned, the number of roots of the equation $\Delta(x, \sigma) = 0$ lying within C_ϵ is equal to the value of the contour integral

$$\vartheta(x) = (1/2\pi i)\int_{C_\epsilon}[\Delta_\sigma(x,\sigma)/\Delta(x,\sigma)]\,d\sigma,$$

where the subscript σ indicates partial derivative. For such values of x the integral $\vartheta(x)$ is obviously continuous in x and has to be a positive integer or zero. Therefore $\vartheta(x) = \vartheta(x_3) = r$ as was to be proved.

For the remaining part of the theorem we show that

$\sigma_\nu(x_4) < \sigma_\nu(x_3)$ for any two numbers x_3 and x_4 such that $x_1 < x_3 < x_4 \leq x_2$. First let η_i, ζ_i be a set of functions identical with the functions $\eta_i^{(1)}(x|x_3)$, $\zeta_i^{(1)}(x|x_3)$ given in (12•6) on the interval x_1x_3 and all zero on x_3x_4. For these functions η_i we have $J_{14}(\eta) = J_{13}(\eta) = \sigma_1(x_3)$. But since ζ_i have discontinuities at x_3, the functions η_i cannot minimize J_{14} and hence $\sigma_1(x_4) < \sigma_1(x_3)$. Next suppose that $\sigma_\nu(x_4) < \sigma_\nu(x_3)$ holds for $\nu \leq \ell$. Let η_i^α, ζ_i^α, $(\alpha = 1, \cdots, \ell+1)$ be sets of functions identical with the functions $\eta_i^\alpha(x|x_3)$, $\zeta_i^\alpha(x|x_3)$ given in (12•6) on the interval x_1x_3 and all zero on x_3x_4. Let $\eta_i = c_\alpha \eta_i^\alpha$, $\zeta_i = c_\alpha \zeta_i^\alpha$ be any linear combination of them with constant coefficients c_α. These constants can always be chosen so that η_i satisfy the norming conditions $N_{14}(\eta) = N_{13}(\eta) = c_\alpha c_\alpha = 1$ and are orthogonal to the functions $\eta_i^\beta(x|x_4)$, $(\beta = 1, \cdots, \ell)$ given by (12•6) with $x_3 = x_4$. Since the functions η_i, ζ_i are identically zero on x_3x_4 and are on x_1x_3 a continuous solution of the system (12•3) with $\sigma\eta_i$ replaced by

$$g_i = \sum_{\alpha=1}^{\ell+1} c_\alpha \sigma_\alpha(x_3)\, \eta_i^\alpha(x|x_3)\,,$$

a relation analogous to (5•9) and the condition $c_\alpha c_\alpha = 1$ will give

$$J_{14}(\eta) = J_{13}(\eta) = \int_{x_1}^{x_3} g_i \eta_i \, dx = \sum_{\alpha=1}^{\ell+1} c_\alpha^2 \sigma_\alpha(x_3) \leq \sigma_{\ell+1}(x_3)\,.$$

Theorem 10•3 then shows that $\sigma_{\ell+1}(x_4) \leq \sigma_{\ell+1}(x_3)$. The equality sign can hold only when η_i, ζ_i form a characteristic solution of the system (12•3) with η_{i3} replaced by η_{i4}. But since η_i, ζ_i all vanish at x_4, they cannot be such a solution. Consequently we must have $\sigma_{\ell+1}(x_4) < \sigma_{\ell+1}(x_3)$. This induction and Lemma 13•1 complete the proof of the theorem.

The following two corollaries are immediate by plotting the curves $\sigma = \sigma_\nu(x_3)$, $(x_1 < x_3 \leq x_2)$ and cutting them by horizontal lines σ = constant.

Corollary 1. Corresponding to a number σ_0 the number of focal points of x_1 which are $< x_2$ is equal to the number of the numbers $\sigma_\nu(x_2)$ which are given by (12·6) with $x_3 = x_2$ and are $< \sigma_0$. The corollary is still true with the inequality sign $<$ throughout replaced by $\leq$. (It is being understood that each focal point is counted with a multiplicity equal to its index).

Corollary 2. If $\sigma' > \sigma''$, then the ν-th focal point of x_1 corresponding to σ'', if it exists, must be preceded by the ν-th focal point of x_1 corresponding to σ', (the order of focal points being counted from left to right).

When the conditions $S_\ell(\eta, \zeta) = 0$ in (12·3) become $\eta_{i1} = 0$, the above theorem and its corollaries remain true with focal points replaced by conjugate points provided the corresponding system is still normal on every subinterval x_1x_3. If we make the stronger assumption that the system

$$\begin{aligned} \eta_i' &= L_{ik}\eta_k + M_{ik}\zeta_k , \\ \zeta_i' &= N_{ik}\eta_k - L_{ki}\zeta_k - \sigma\eta_i , \end{aligned} \qquad \eta_{i3} = \eta_{i4} = 0, \quad (x_1 \leq x_3 < x_4 \leq x_2) \tag{13·6}$$

is normal on every subinterval x_3x_4, then we have the following stronger theorem whose proof is practically the same as those of Lemma 13·1 and Theorem 13·1.

Theorem 13·2. If $\sigma_\nu(x_3, x_4)$ with $\sigma_1(x_3, x_4) \leq \sigma_2(x_3, x_4) \leq \cdots$ represent the totality of characteristic numbers of the system (13·6), then each of the numbers $\sigma_\nu(x_3, x_4)$ varies continuously with x_3, x_4 and properly increases as the

interval x_3x_4 shrinks. Moreover all of them become positively infinite as the interval shrinks to a point.

Corollary. Corresponding to a number σ_0 the number of conjugate points of x_3 between x_3 and x_4 is the same as the number of conjugate points of x_4 between the same limits and is equal to the number of the numbers $\sigma_\nu(x_3, x_4)$ which are $<$ or $\not> \sigma_0$ according as the word "between" is used in its strict or broader sense. Moreover, when σ_0 remains fixed and x_3 varies, each conjugate point of x_3, so long as it remains between x_1 and x_2 in the strict sense, varies continuously with x_3 and advances or regresses with it.

The first sentence of the corollary is obvious. As to the second a little justification is required. To do this let us understand temporarily the conjugate points of a given point as its conjugate points corresponding to σ_0. This being so, let x_3 have an initial position t such that its ν-th right conjugate point $\bar{t}$ exists and is to the left of x_2. As in the

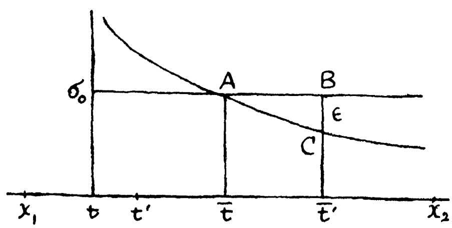


figure draw a part of the curve $\sigma = \sigma_\nu(t, x)$ and let it meet the line $\sigma = \sigma_0$ at a point A. The projection of A on the x-axis is of course $\bar{t}$. Let B be any point on the line $\sigma = \sigma_0$ and be near and to the right of A. Draw the vertical line $B\bar{t}'$ to meet the curve at a point C. Let ϵ be the length of BC. By the continuity and monotonicity of the numbers

$\sigma_\nu(x_3, x_4)$, it follows that there exists a point t' such that whenever a point t'' is $> t$ and $< t'$ we must have

$$0 < \sigma_\nu(t'', \bar{t}') - \sigma_\nu(t, \bar{t}') < \epsilon .$$

This shows that the curve $\sigma = \sigma_\nu(t'', x)$ must meet the segment BC and therefore the segment AB at some interior points; i.e. the ν-th right conjugate point $\bar{t}''$ of t'' exists and is $> \bar{t}$ and $< \bar{t}'$. As B may be chosen arbitrarily near A, this completes the proof for the above situation. The proof for other cases is the same.

14. <u>First comparison theorems</u>. We begin by studying the effect of an increase or decrease of the quadratic form $2q(\eta)$ on the numbers $\sigma_\nu(x_3)$ and consequently on the focal points of x_1. At the same time we allow a change of the end conditions $\Psi_\mu = 0$. To make the statement clear we let

$$\bar{J}_{13}(\eta) = 2\bar{q}(\eta) + \int_{x_1}^{x_3} 2\omega(x, \eta, \eta')\,dx, \text{ where } 2\bar{q}(\eta) = \bar{A}_{ik}\eta_{i1}\eta_{k1}, \tag{14·1}$$

and compare the problem of minimizing $J_{13}(\eta)$ in (12·4) under the conditions (12·5) with the problem of minimizing $\bar{J}_{13}(\eta)$ under the conditions

$$\Phi_\alpha(x, \eta, \eta') = 0, \quad N_{13}(\eta) = 1, \quad \bar{\Psi}_\kappa(\eta) = \bar{a}_{\kappa i}\eta_{i1} = 0, \quad \eta_{i3} = 0. \tag{14·2}$$

The boundary value system for the latter problem

$$\begin{aligned} &\eta_i' = L_{ik}\eta_k + M_{ik}\zeta_k, && \bar{S}_\ell(\eta, \zeta) = \bar{a}_{\ell i}\eta_{i1} - \bar{b}_{\ell i}\zeta_{i1} = 0, \\ &\zeta_i' = N_{ik}\eta_k - L_{ki}\zeta_k - \sigma\eta_i, && \eta_{i3} = 0, \end{aligned} \tag{14·3}$$

differs from (12·3) only in the conditions $\bar{S}_\gamma = 0$. We shall assume that this system is also normal on every subinterval x_1x_3 and analogously to (12·6) let

$$\left\{ \bar{\sigma}_\nu(x_3),\ \bar{\eta}_i^\nu(x|x_3),\ \bar{\zeta}_i^\nu(x|x_3) \right\} \text{ with } \bar{\sigma}_1(x_3) \le \bar{\sigma}_2(x_3) \le \cdots \tag{14.4}$$

be the totality of its characteristic numbers and the corresponding orthonormalized characteristic solutions.

Theorem 14·1. Let t be the number of the conditions $\bar{\Psi}_\tau = 0$ which are linearly independent of $\Psi_\mu = 0$, and r be the number of conditions which have to be adjoined to $b_{\chi 1}\zeta_{i1} = \bar{b}_{\chi 1}\zeta_{i1} = 0$ in order to form n linearly independent conditions. C_1) If $q(\eta) \ge \bar{q}(\eta)$ with respect to all sets of constants η_{i1} not all zero and satisfying the conditions $\Psi_\mu = \bar{\Psi}_\tau = 0$, then $\sigma_{\nu+t}(x_3) \ge \bar{\sigma}_\nu(x_3)$ for all values of x_3 and ν. C_2) If the strengthened condition $q(\eta) > \bar{q}(\eta)$ holds under the above restrictions, then not only $\sigma_{\nu+t}(x_3) \ge \bar{\sigma}_\nu(x_3)$ but also surely $\sigma_{\nu+t+r}(x_3) > \bar{\sigma}_\nu(x_3)$ for all values of x_3 and ν.

We shall prove C_2) only because the proof of C_1) is similar and easier. First we select r sets of constants $d_{\gamma i}$, ($\gamma = 1,\cdots,r$) such that $d_{\gamma i}\zeta_{i1} = b_{\chi 1}\zeta_{i1} = \bar{b}_{\chi 1}\zeta_{i1} = 0$ have no solution except $\zeta_{i1} = 0$. For any fixed value of x_3 and ν, let

$$\eta_i = c_\alpha \eta_i^\alpha(x|x_3),\quad \zeta_i = c_\alpha \zeta_i^\alpha(x|x_3),\quad (\alpha = 1,\cdots,\nu+t+r)$$

be a linear combination of the functions in (12·6). Constants c_α can always be chosen so that ζ_i satisfy the conditions $d_{\gamma i}\zeta_{i1} = 0$ while η_i satisfy the conditions $N_{13}(\eta) = c_\alpha c_\alpha = 1$, $\bar{\Psi}_\tau = 0$ and are, if $\nu > 1$, orthogonal to the functions $\eta_i^\beta(x|x_3)$, ($\beta = 1,\cdots,\nu-1$) in (14·4). Since the functions η_i, ζ_i form a solution of the system (12·3) with $\sigma\eta_i$ replaced by

$$g_i = \sum_{\alpha=1}^{\nu+t+r} c_\alpha \sigma_\alpha(x_3)\eta_i^\alpha(x|x_3),$$

we have by a formula like (5•9) and the relation $c_\alpha c_\alpha = 1$

$$J_{13}(\eta) = \int_{x_1}^{x_3} g_i \eta_i dx = \sum_{\alpha=1}^{\nu+t+r} c_\alpha^2 \sigma_\alpha(x_3) \leq \sigma_{\nu+t+r}(x_3).$$

This and the hypothesis in C_2) give $J_{13}(\eta) \leq \sigma_{\nu+t+r}(x_3)$ in which the equality sign holds only when η_{i1} are all zero. Hence, according to Theorem 10•3, $\bar{\sigma}_\nu(x_3) < \sigma_{\nu+t+r}(x_3)$ unless the end values η_{i1} are all zero and η_i, ζ_i are a characteristic solution of (14•3). But η_i, ζ_i cannot be such a solution as is seen by a consideration of the conditions $S_\ell(\eta,\zeta) = \bar{S}_\ell(\eta,\zeta) = 0$ and $d_{\gamma i}\zeta_{i1} = 0$. The inequality is thus established.

In terms of focal points we have the following

Corollary. Under the hypothesis in C_1) and corresponding to a number σ_0, the $(\nu + t)$-th focal point of x_1 relative to (12•3), if it exists, is not on the left of the ν-th focal point of x_1 relative to (14•3). Under the hypothesis in C_2) and corresponding to σ_0, the $(\nu + t + r)$-th focal point of x_1 relative to (12•3), if it exists, must follow the ν-th focal point of x_1 relative to (14•3).

As an application of this corollary we easily deduce the following theorem which was used in the proof of Lemma 11•3.

Theorem 14•2. Let (D) represent the differential equations in (14•3) with $\sigma = 0$. Assume that (D) is normal on every subinterval x_1x_3 and there is no conjugate point of x_1 relative to (D) on $x_1 < x \leq x_2$. Then there exists a conjugate system of solutions η_i^ℓ, ζ_i^ℓ, $(\ell = 1, \cdots, n)$ of (D) such that the determinant $|\eta_i^\ell(x)|$ is everywhere different from zero on $x_1 \leq x \leq x_2$.

Let $\bar{\eta}_i^\ell$, $\bar{\zeta}_i^\ell$, $(\ell = 1, \cdots, n)$ be any n sets of

independent solutions of (D) with initial values $\bar{\eta}_i^l(x_1) = 0$ so that they form a conjugate system of solutions. By means of the hypothesis in the theorem it follows that $|\bar{\eta}_i^l(x)| \neq 0$ on $x_1 < x \leq x_2$. We may therefore suppose that these solutions have been normed so that $\bar{\eta}_i^l(x_2) = \delta_{il}$, $\bar{\zeta}_i^l(x_2) = -g_{i\gamma}$, where δ_{il} is the Kronecker δ and $g_{i\gamma}$ is some constant. By the conjugate property of the system we must have $g_{i\gamma} = g_{\gamma i}$. Now let η_i^l, ζ_i^l be n sets of solutions of (D) with end values $\eta_i^l(x_2) = \delta_{il}$, $\zeta_i^l(x_2) = -\delta_{il} - g_{i\gamma}$. Evidently these solutions also form a conjugate system. I say the determinant $|\eta_i^l(x)|$ is everywhere different from zero on $x_1 \leq x \leq x_2$. Indeed let

$$J_{32}(\eta) = 2q(\eta) + \int_{x_3}^{x_2} 2\omega(x,\eta,\eta')dx, \quad 2q(\eta) = (\delta_{ik} + g_{ik})\eta_{i2}\eta_{k2},$$
$$\bar{J}_{32}(\eta) = 2\bar{q}(\eta) + \int_{x_3}^{x_2} 2\omega(x,\eta,\eta')dx, \quad 2\bar{q}(\eta) = g_{ik}\eta_{i2}\eta_{k2},$$

and consider the problems of minimizing J_{32} and $\bar{J}_{32}$ with respect to all functions $\eta_i(x)$ of class D' on x_3x_2, $(x_1 \leq x_3 < x_2)$ and satisfying the conditions $\Phi_\alpha(x, \eta, \eta') = 0$, $N_{32}(\eta) = 1$, $\eta_{i3} = 0$. No condition is put at x_2. Their boundary value problems are therefore normal on every subinterval x_3x_2. We can easily verify that η_i^l, ζ_i^l and $\bar{\eta}_i^l$, $\bar{\zeta}_i^l$ satisfy the respective boundary conditions at x_2. By the relation $|\bar{\eta}_i^l(x)| \neq 0$ on $x_1 < x \leq x_2$, we see that the first focal point of x_2 relative to the problem of minimizing $\bar{J}_{32}$ and corresponding to $\sigma = 0$ is at x_1. By the corollary to Theorem 14·1 with x_2, x_3 in place of x_1, x_3 and $t = r = 0$, we see that the first focal point of x_2 relative to the problem of minimizing J_{32} and corresponding to $\sigma = 0$ cannot be on the interval $x_1 \leq x < x_2$. Since $|\eta_i^l(x_2)| = 1$, the non-vanishing of $|\eta_i^l(x)|$ on $x_1 \leq x$

$\leq x_2$ is therefore proved.

Next we shall make a change in the integrand function $2\omega(x, \eta, \eta')$ as well as in the quadratic form $2q(\eta)$ and the end conditions $\Psi_\mu = 0$ but hold the differential conditions $\Phi_\alpha(x, \eta, \eta') = 0$ unchanged. To be more specific let us compare the problem of minimizing $J_{13}(\eta)$ in (12•4) under the conditions (12•5) with the problem of minimizing another similar expression $\tilde{J}_{13}(\eta)$ under the conditions

$$\Phi_\alpha(x,\eta,\eta') = 0, \quad N_{13}(\eta) = 1, \quad \tilde{\Psi}_c(\eta) = \tilde{a}_{ci}\eta_{i1} = 0, \quad \eta_{i3} = 0. \tag{14•5}$$

We suppose that all the fundamental hypotheses corresponding to those satisfied by the first problem are satisfied by the second. The boundary value system for the second problem may differ from (12•3) both in the coefficients L_{ik}, M_{ik}, N_{ik} and in the conditions $S_\gamma = 0$. By analogy with (12•6) let $\tilde{\sigma}_\nu(x_3)$ with $\tilde{\sigma}_1(x_3) \leq \tilde{\sigma}_2(x_3) \leq \cdots$ be the characteristic numbers for the new system. We have the following theorem whose proof is similar to that of Theorem 14•1 and may be omitted.

<u>Theorem</u> 14•3. <u>Let</u> t <u>be the number of the conditions</u> $\tilde{\Psi}_c = 0$ <u>which are linearly independent of</u> $\Psi_\mu = 0$. <u>If, on every subinterval</u> x_1x_3, <u>we have</u> $J_{13}(\eta) > \tilde{J}_{13}(\eta)$ <u>for all functions</u> $\eta_i(x)$ <u>satisfying the conditions</u> (12•5) <u>and</u> (14•5), <u>then</u> $\sigma_{\nu+t}(x_3) > \tilde{\sigma}_\nu(x_3)$ <u>for all values of</u> x_3 <u>and</u> ν.

<u>Corollary. Under the hypothesis in the theorem and corresponding to a number</u> σ_0, <u>the</u> $(\nu + t)$-th <u>focal point of</u> x_1 <u>relative to the problem of minimizing</u> J_{13}, <u>if it exists, must follow the</u> ν-th <u>focal point of</u> x_1 <u>relative to the problem of minimizing</u> $\tilde{J}_{13}$.

15. Further comparison theorems. In this section we shall deduce some comparison theorems which are different in character from those of the preceding section. We begin by comparing the general system (5·1) with the special system

$$\begin{aligned} \eta_i' &= L_{ik}\eta_k + M_{ik}\zeta_k , \\ \zeta_i' &= N_{ik}\eta_k - L_{ki}\zeta_k - \sigma\eta_i , \end{aligned} \qquad \eta_{i1} = \eta_{i2} = 0 . \tag{15·1}$$

Note that this is the system arising from minimizing $J^0(\eta) = \int_{x_1}^{x_2} 2\omega(x, \eta, \eta')dx$ with respect to all functions satisfying the conditions $\Phi_\alpha(x, \eta, \eta') = 0$, $N(\eta) = 1$, $\eta_{i1} = \eta_{i2} = 0$. If we assume that this system is normal on x_1x_2, then we have the following

Theorem 15·1. C_1) The number of characteristic numbers of the system (5·1) less than a number σ_0 lies between χ and $\chi + 2n - p$ inclusive, where χ is the number of characteristic numbers $< \sigma_0$ of the system (15·1) and p is the number of the end conditions $\Psi_\mu(\eta) = 0$ in Problem 2 or, what is the equivalent thing, $2n - p$ is the rank of the matrix formed by the coefficients of ζ_{i1}, ζ_{i2} in the boundary conditions $S_\rho = 0$ of (5·1). C_2) The preceding statement is still true with $< \sigma$ replaced by $\leq \sigma_0$.

In order to prove C_1) let $\sigma_1 \leq \sigma_2 \leq \cdots \leq \sigma_h$ be all the characteristic numbers of (5·1) which are $< \sigma_0$ and

$$\eta_i^\alpha(x), \quad \zeta_i^\alpha(x), \quad (\alpha = 1, \cdots h) \tag{15·2}$$

be the corresponding orthonormal characteristic solutions of (5·1). Similarly let $\bar\sigma_1 \leq \bar\sigma_2 \leq \cdots \leq \bar\sigma_l$ be all the characteristic numbers of (15·1 which are $< \sigma_0$ and

$$\bar\eta_i^\beta(x), \quad \bar\zeta_i^\beta(x), \quad (\beta = 1, \cdots, l) \tag{15·3}$$

be the corresponding orthonormal characteristic solutions of (15·1). The notations σ_{k+1} and $\bar{\sigma}_{l+1}$ will denote the first characteristic numbers of (5·1) and (15·1) respectively which are $\geq \sigma_0$.

First we prove that $h \ngtr \gamma + 2n - p$. For, any linear combination $\eta_i = c_\alpha \eta_i^\alpha$, $\zeta_i = c_\alpha \zeta_i^\alpha$ with constant coefficients c_α of the functions in (15·2) is a solution of the system (5·1) with $\sigma\eta_i$ replaced by

$$(15\cdot4) \qquad g_i = \sum_{\alpha=1}^{h} c_\alpha \sigma_\alpha \eta_i^\alpha(x);$$

and since the boundary conditions $S_\rho = 0$ in (5·1) imply the end conditions $\Psi_\mu(\eta) = 0$, it amounts to $2n - p$ conditions on $\eta_i = c_\alpha \eta_i^\alpha$ to require the conditions $\eta_{i1} = \eta_{i2} = 0$. Consequently, if $h > \gamma + 2n - p$, constants c_α could be determined so that η_i would satisfy the conditions $N(\eta) = c_\alpha c_\alpha = 1$, $\eta_{i1} = \eta_{i2} = 0$ and would be orthogonal to the functions $\bar{\eta}_i^\beta$ in (15·3). For these functions η_i, Theorem 10·3 would require the expression $J^0(\eta)$ to be $\geq \bar{\sigma}_{l+1}$. But (5·8), (15·4) and the condition $c_\alpha c_\alpha = 1$ would give

$$J^0(\eta) = \int_{x_1}^{x_2} g_i \eta_i \, dx = \sum_{\alpha=1}^{h} c_\alpha^2 \sigma_\alpha \leq \bar{\sigma}_{l+1}.$$

This contradiction shows that $h \ngtr \gamma + 2n - p$.

Next we prove that $h \nless \gamma$. For any linear combination $\bar{\eta}_i = d_\beta \bar{\eta}_i^\beta$, $\bar{\zeta}_i = d_\beta \bar{\zeta}_i^\beta$ with constant coefficients d_β of the functions in (15·3) is a solution of the system (15·1) with $\sigma\eta_i$ replaced by

$$(15\cdot5) \qquad \bar{g}_i = \sum_{\beta=1}^{l} d_\beta \bar{\sigma}_\beta \bar{\eta}_i^\beta(x).$$

If $h < \gamma$, constants d_β could be chosen so that $\bar{\eta}_i$ would satisfy

the condition $N(\eta) = d_\beta d_\beta = 1$ and would be orthogonal to the functions η_i^α in (15·2). For these functions $\bar{\eta}_i$, Theorem 10·3 would require the expression $J(\eta)$ in Problem 2 to be $\geq \sigma_{k+1}$. But (5·8), (15·5) and the condition $d_\beta d_\beta = 1$ would give

$$J(\bar{\eta}) = \int_{x_1}^{x_2} \bar{g}_i \bar{\eta}_i \, dx = \sum_{\beta=1}^{l} d_\beta^2 \bar{\sigma}_\beta < \sigma_0 \leq \sigma_{k+1} .$$

This contradiction shows that $h \not< \gamma$.

The proof of C_2) is exactly the same by supposing that $\sigma_1 \leq \sigma_2 \leq \cdots \cdots \leq \sigma_k \leq \sigma_0 < \sigma_{k+1}$ and $\bar{\sigma}_1 \leq \bar{\sigma}_2 \leq \cdots \cdots \leq \bar{\sigma}_l \leq \sigma_0 < \bar{\sigma}_{l+1}$.

The following corollary is immediate by suitably applying C_1) or C_2) of the theorem to two numbers σ' and σ''.

Corollary 1. The number of characteristic numbers of (5·1) on any finite interval (σ', σ''), open or closed or semi-closed at pleasure, differs from the corresponding number of (15·1) by at most $2n - p$.

In terms of conjugate points the above theorem and Corollary 1 to Theorem 13·1 gives the following corollary, provided (15·1) is also normal on every subinterval x_1x_3.

Corollary 2. The number of characteristic numbers of (5·1) lies between γ and $\gamma + 2n - p$ inclusive, where γ is the number of conjugate points of x_1 which are $< x_2$ and corresponding to σ_0 and p has the same meaning as in the theorem. The corollary is still true with the inequality sign $<$ throughout replaced by $\leq$.

Next we have the following

Theorem 15·2. C_1) When the boundary conditions in the system (5·1) change from $S_\rho = 0$ to $\bar{S}_\rho = 0$, the number of characteristic numbers $<$ or $\not>$ a number σ_0 can change by at most the greater of the two numbers $2n - p$ and $2n - \bar{p}$, where

<u>p has the same meaning as in Theorem 15·1 and $\bar{p}$ is the number corresponding to p for the conditions $\bar{S}_\rho = 0$.</u> C_2) <u>If the rank of the matrix of the coefficients of $S_\rho = 0$, $\bar{S}_\rho = 0$ is $4n - r$, then the above change of number must also be $\leq 2n - r$.</u>

The statement C_1) is a corollary of the preceding theorem. The proof of C_2) is made by arguments similar to those used in the preceding proof and hence will be omitted.

When we apply the above theorem to two systems of the form (12·3) which differ only in the conditions $S_\gamma = 0$ and interpret the results in terms of focal points we have the following corollary which is a generalization of the well-known Sturm's separation theorem.

<u>Corollary.</u> <u>Corresponding to a fixed number σ_0, the number of focal points of x_1 on a given interval ($x_1 < x \leq x_3$) or ($x_1 < x < x_3$) relative to two conjugate families can differ by at most n. If the two families have r linearly independent members in common, the difference is at most $n - r$.</u>

JACOBI'S CONDITION FOR MULTIPLE INTEGRAL PROBLEMS OF THE CALCULUS OF VARIATIONS

BY

ALBERT WILLIAM RAAB

TABLE OF CONTENTS

INTRODUCTION

The necessary condition of Jacobi, which limits the range within which an extremal surface may give a double integral a minimum was treated by Kobb (5)[1] in 1892 for the parametric integral

$$I = \iint \Phi(x, y, z, x_u, y_u, z_u, x_v, y_v, z_v)\,du\,dv.$$

At best Kobb's method is complicated and, in the present state of the theory of partial differential equations, does not lend itself to a proof of Jacobi's condition for the more general problems considered in this paper. Sommerfeld (6), extending a method due to Schwarz for simple integrals, in 1899 proved Jacobi's condition for the non-parametric integral

$$I = \iint f(x, y, z, p, q)\,dx\,dy. \tag{1}$$

More recently several papers have appeared (7 to 15) which treat the condition for the latter integral and for the integral

$$I = \iint f(x, y, z, p, q, r, s, t)\,dx\,dy \tag{2}$$

with and without the respective isoperimetric conditions

$$K = \iint g(x, y, z, p, q)\,dx\,dy = \chi \tag{1'}$$

and

$$K = \iint g(x, y, z, p, q, r, s, t)\,dx\,dy = \chi. \tag{2'}$$

For the integral (1) an interesting advance has been made

[1] Numbers in parentheses refer to bibliography on page 24.

by Haar (15), who, in pursuing the consequences of his well-known fundamental lemma, has succeeded in replacing the second order partial differential equations of Lagrange and of Jacobi by two systems of first order equations, thus dispensing with the customary hypothesis that the extremal surface be of class C". Noteworthy results have also been attained by Lichtenstein (11, 12) and Picone (13) who state Jacobi's condition in terms of characteristic numbers of a boundary value problem, reminiscent of Hilbert's treatment of Jacobi's condition for simple integrals. A somewhat analogous statement of Jacobi's condition is given by Simmons (14) for the integral (1) with variable boundary.

But Jacobi's condition for the more general problems involving multiple integrals appears to have received hitherto little or no attention. In this paper we prove the analogues of Jacobi's condition for both the non-parametric integral

$$(3)\quad I = \int_S^n f(x_1,\cdots,x_n;\ y_1,\cdots,y_m;\ \partial y_1/\partial x_1,\cdots,\ \partial y_m/\partial x_n)dx_1\cdots dx_n$$

and the parametric integral

$$(4)\qquad I = \int_S^n F(y_1,\ \cdots,\ y_m;\ \partial y_1/\partial u_1,\ \cdots,\ \partial y_m/\partial u_n)du_1\cdots du_n,$$

the boundaries being fixed. Incidentally several lemmas are proved which aid in analyzing the multiple integral problem.

The method here employed is one that originated with Bliss[1] and which, because of its elegance, has been applied by him and his pupils to various problems involving simple integrals and by Kubota (9) to the double integral (2) with and without the side condition (2'). One of the chief points of interest of

---

[1] Bliss, Jacobi's Conditions for Problems of the Calculus of Variations in Parametric Form, _Transactions of the American Mathematical Society_, vol. 17, (1916), p. 195.

this paper is the simplicity attending the application of this method to the general parametric problem (4). Attempts to treat this problem by the methods of authors cited in preceding paragraphs lead to great complications if not, indeed, to insurmountable difficulties.

The statement of Jacobi's condition given in parts I and II is couched in terms of a solution of the Jacobi equations whose first partial derivatives are not all identically zero along the manifold on which the solution vanishes. For simple special cases it is well known that a solution of an elliptic partial differential equation (such as the Jacobi equation for our problem) which vanishes together with its first partial derivatives on the boundary of a region, vanishes identically on the interior of the region. In part III is given an outline of the method of proof for the last named result and a survey of the literature relating thereto for the general case of a single linear homogeneous elliptic equation of the second order in n independent variables. Reference is also made to extensions of portions of this theory to certain special cases of systems of equations.

PART I

THE ANALOGUE OF JACOBI'S CONDITION FOR THE NON-PARAMETRIC PROBLEM

1. Notations and definitions. The symbols t, x, y will stand for points in the spaces T, X, Y, of n - 1, n, and m dimensions respectively. Repeated subscripts will indicate summations, the ranges being

$$\gamma = 1, 2, \cdots, n - 1; \alpha, \beta = 1, 2, \cdots, n;$$
$$i, k = 1, 2, \cdots, m.$$

The matrix $||p_i^\alpha||$, where $p_i^\alpha = \partial y_i/\partial x_\alpha$ will be represented by p, and a set of equations such as $x_\alpha = g_\alpha(t_1, \cdots, t_{n-1})$, $\alpha = 1, 2, \cdots, n$ will be represented by the single equation $x = g(t)$. Further, the symbols S, M, L, W, have the meanings:

S $\equiv$ a bounded connected open point set in n-dimensional space;

M $\equiv$ the boundary of S; M is assumed to have zero n-dimensional Jordan content;

L $\equiv$ a bounded (n - 1)-dimensional manifold in XY-space whose projection on X-space is M, the correspondence between the points of L and M being one-to-one;

W = a region of the XYP-space.

Definition I. A simple regular n-dimensional manifold $y = y(x)$ with boundary L is one such that the functions $y = y(x)$ are of class C' in S + M and such that for each point x of M the point $(x, y(x))$ is in L.

Definition II. A regular n-manifold $y = y(x)$ is one such that

a. It consists of a finite nūmber of fragments each of which is a simple regular manifold.

b. The functions $y = y(x)$ are continuous and single-valued in S.

Definition III. An admissible manifold is one which is regular and has its elements intérior to the region W.

2. The Lagrange equations and the edge conditions. For multiple integrals in the calculus of variations we have the

Problem: To find in a given class of admissible manifolds

$$y = y(x), \qquad (x \text{ in } S),$$

which pass through the boundary manifold L, one which minimizes the m-tuple integral

$$I = \int_S f(x,\, y,\, p)dx \tag{2:1}$$

where f is of class C'' in the region W.

If for this problem $\mathfrak{M}$ is a minimizing manifold $y = y(x)$, there follows the[1]

THEOREM I. On every part of $\mathfrak{M}$ where the functions $y_i(x)$ have continuous second derivatives, we have

$$\partial f_i^{\alpha} / \partial x_{\alpha} - f_{y_i} = 0, \qquad f_i^{\alpha} \equiv f_{p_i^{\alpha}}.$$

Definition IV. A solution $y = y(x)$ of the above problem will be said to be of class D'' if the functions $y = y(x)$ define a regular n-dimensional manifold in the interior of each cell of which they possess continuous second partial derivatives, while

[1] Powell, Edge Conditions for Multiple Integrals, Chicago Dissertation, 1931.

at some points on the common boundary of two cells some of the first partial derivatives are discontinuous.

Let E' be an (n- 1)-dimensional manifold which is a common boundary of two of the cells of $\mathfrak{M}$. Consider a part R of E' whose projection E on X-space is cut by a parallel to an axis in at most a finite number of points and segments and which can be represented by the equations

$$x = g(t), \qquad (t \text{ in } T),$$

where the functions g are of class C' and $\sum A_\alpha^2 \neq 0$ for t in T,

$$A_\alpha \equiv (-1)^{\alpha-1} \frac{\partial(g_1, \cdots, g_{\alpha-1}, g_{\alpha+1}, \cdots, g_n)}{\partial(t_1, \cdots, t_{n-1})}. \tag{2:2}$$

For what we have called a solution of class D", the following necessary conditions must be satisfied.[1]

THEOREM II. <u>If</u> $y = y(x)$ <u>define a minimizing manifold of class</u> D", <u>then at points of</u> R <u>the expressions</u> $f_i^\alpha A_\alpha$ <u>must be continuous, that is,</u>

$$f_i^\alpha A_\alpha \Big|^1 = f_i^\alpha A_\alpha \Big|^2, \qquad (i = 1, \cdots, m),$$

<u>the vertical bars indicating that the limiting values of the functions</u> $f_i^\alpha A_\alpha$ <u>are obtained by approaching a point of</u> R <u>first from one side of</u> R <u>and then from the other side.</u>

3. <u>The second variation and the Jacobi equations.</u> If $y = y(x)$ minimize

$$I = \int_S f(x, y, p)dx$$

we find, as usual, by considering a one-parameter family of ad-

---

[1] Ibid. pp. 46 ff.

missible manifolds

$$y = y(x) + \varepsilon\eta(x), \qquad \eta = 0 \text{ on } M,$$

that the relations $I'(0) = 0$ and $I''(0) \geqq 0$ must hold. Here $I''(0)$, the second variation, equals

$$\int_S 2\Omega(x, \eta, q)dx$$

where 2Ω is a quadratic form in the arguments $\eta = (\eta_1, \cdots, \eta_m)$ and $q_i^\alpha = \partial\eta_i/\partial x_\alpha$. The coefficients of this form are the second derivatives of f with respect to its arguments y, p.

<u>Definition V.</u> An (n - 1)-dimensional manifold $\overline{M}$ will be said to be a Green's manifold if it is closed and bounds a portion $\overline{S}$ of n-dimensional X-space such that the formula holds

$$\int_{\overline{S}}^{n} \sum (\partial P_\alpha/\partial x_\alpha)dX = -\int_{\overline{M}}^{n-1} \sum P_\alpha \cos(n, x_\alpha)d\overline{M},$$

where the P_α are of class C' in $\overline{S} + \overline{M}$ and the $\cos(n, x_\alpha)$ are the direction cosines of the inner normal of $\overline{M}$.

We find this definition convenient because the degree of generality of the boundary manifold $\overline{M}$ in the various extensions of Green's Theorem depends upon the value of n. Thus for n = 2 one may not merely dispense with the customary restriction that $\overline{M}$ be met by a line parallel to an axis in only a finite number of points and segments, but may assume $\overline{M}$ to be any simply closed rectifiable curve.[1] A somewhat analogous theorem holds for n = 3.[2] An (n - 1)-dimensional manifold will surely be a Green's

---

[1] Grosz, Monatshefte fur Mathematik und Physik, vol. 27 (1916), p. 70.

[2] Lichtenstein, Archiv der Mathematik und Physik, 1918, p. 31. For Green's Theorem in n-dimensions, see Brouwer, Proceedings of Amsterdam Academy, 1919; Franklin, Annals of Mathematics, 1923, p. 213; Giraud, Annales Ecole Norm., 1926, p. 1.

manifold if the region that it bounds can be broken up into a finite number of sub-regions, each of which is bounded by a manifold of spherical type that is met by a line parallel to an axis in only a finite number of points and segments.

If the minimizing manifold y = y(x) and the admissible variations $\eta(x)$ are of class C" and the boundary M is a Green's manifold, we may transform the second variation as follows. By Euler's Theorem on homogeneous functions

$$2\Omega = \Omega_{\eta_i}\eta_i + \Omega_i^\alpha q_i^\alpha \ , \quad \Omega_i^\alpha \equiv \Omega_{q_i^\alpha}$$

Also

$$\frac{\partial(\Omega_i^\alpha \eta_i)}{\partial x_\alpha} = \Omega_i^\alpha q_i^\alpha + \eta_i \frac{\partial \Omega_i^\alpha}{\partial x_\alpha} \ ,$$

so that Green's Theorem enables us to write

$$\begin{aligned} I''(0) &= \int_S \left[\eta_i\left(\Omega_{\eta_i} - \frac{\partial \Omega_i^\alpha}{\partial x_\alpha}\right) + \frac{\partial(\Omega_i^\alpha \eta_i)}{\partial x_\alpha}\right] dx \\ &= \int_S \eta_i\left(\Omega_{\eta_i} - \frac{\partial \Omega_i^\alpha}{\partial x_\alpha}\right) dx \end{aligned}$$

since the η's vanish on M.

The m equations

$$\Omega_{\eta_i} - \frac{\partial \Omega_i^\alpha}{\partial x_\alpha} = 0 \qquad (i = 1,2,\cdots,m)$$

are the analogues of the Jacobi equations for problems involving only simple integrals. It is evident that they are of the second order and linear and homogeneous in the η's and their first and second derivatives. The symmetry possessed by the system is more apparent when it is written in the form

$$\begin{aligned} \Delta_{11}\eta_1 + \cdots + \Delta_{1m}\eta_m + \Psi_1 &= 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ \Delta_{m1}\eta_1 + \cdots + \Delta_{mm}\eta_m + \Psi_m &= 0 \end{aligned} \tag{3:1}$$

Here Δ is the operator

$$\Delta_{ik} = \Delta_{ki} = f_{ik}^{\alpha\beta}\,(\partial^2/\partial x_\alpha \partial x_\beta), \quad f_{ik}^{\alpha\beta} \equiv f_{p_i^\alpha\, p_k^\beta},$$

and the ψ's are linear functions of the η's and their first partial derivatives. We proceed to examine this system more closely, assuming, as always, that the functions $y = y(x)$ minimize (2:1).

<u>Definition</u> <u>VI</u>. By Legendre's condition in the stronger form is meant that along the extremizing manifold the quadratic form

$$f_{ik}^{\alpha\beta}\, h_\alpha^i h_\beta^k \tag{3:2}$$

is different from zero for all matrices h_α^i of rank one, that is, for all non-zero sets of collinear m-dimensional vectors

$$h_\alpha^1, \cdots, h_\alpha^m = c_\alpha v_1, \cdots, c_\alpha v_m \qquad (\alpha = 1, \cdots, n),$$

being positive for a minimum and negative for a maximum.[1]

LEMMA I. <u>Legendre's condition in the stronger form is equivalent to the condition that along the extremizing manifold the quadratic form</u>

$$B_{ik} v_i v_k, \quad \underline{\text{where}} \quad B_{ik} = f_{ik}^{\alpha\beta}\, c_\alpha c_\beta \tag{3:3}$$

<u>be definite for every non-zero set</u> $c_1, \cdots, c_n$.

Let v_i be any non-zero vector and c_α any non-zero set of constants. Substituting $h_\alpha^i = c_\alpha v_i$ in (3:3), we obtain

$$f_{ik}^{\alpha\beta}\, h_\alpha^i h_\beta^k = f_{ik}^{\alpha\beta}\, c_\alpha c_\beta v_i v_k = B_{ik} v_i v_k, \quad B_{ik} = f_{ik}^{\alpha\beta}\, c_\alpha c_\beta,$$

a definite form.

[1] For the necessity of Legendre's condition see (1,2,3,4).

Note that B_{ik} is precisely what we obtain if in the differential operator Δ_{ik} we replace $(\partial^2/\partial x_\alpha \partial x_\beta)$ by $c_\alpha c_\beta$. Therefore, the discriminant of (3:3) is the characteristic form of the system (3:1). Moreover, since (3:3) is a definite form, its discriminant, and therefore the characteristic form of (3:1), are different from zero for all non-zero sets c_α. That is, the determinant

$$|f_{ik}^{\alpha\beta} c_\alpha c_\beta| \neq 0 \quad \text{for } (c_1, \cdots, c_n) \neq (0, \cdots, 0);$$

so that, by analogy with the theory of the single second order partial differential equation

$$(3:4) \qquad a_{\alpha\beta}(x)(\partial^2 u/\partial x_\alpha \partial x_\beta) + b_\alpha(x)(\partial u/\partial x_\alpha) + c(x)u + d(x) = 0$$

we might characterize the system (3:1) as elliptic in the region S. When $m = 1$, this system reduces to an elliptic equation of the form (3:4) with $d(x) \equiv 0$.

4. The analogue of Jacobi's condition. In this section we shall prove the

THEOREM III.

Hypotheses:

1. The functions $y = y(x)$ define a minimizing manifold of class C" on which Legendre's condition in the stronger form is satisfied.

2. $\overline{M}$ is any $(n - 1)$-dimensional Green's manifold bounding a domain $\overline{S}$ such that $\overline{S}$ is in the domain of integration S of X-space and the point set $S - (\overline{S} + \overline{M})$ consists of a finite number of sets of type S.

Conclusion:

There cannot exist a solution u of the system of Jacobi

equations which is of class C'' on $\bar{S} + \bar{M}$, vanishes on $\bar{M}$ and is such that the first partial derivatives are not all zero at a point P of $\bar{M}$ in S.

To prove the theorem, suppose a solution u such as described existed with, say, $\partial u_1/\partial x_n \neq 0$, at the point P. Then we know that the portion E of $\bar{M}$ in the neighborhood of P can be represented by the equation $x_n = x_n(x_1, \cdots, x_{n-1})$ where x_n is a single-valued function of its arguments of class C''. It will be assumed that near P, E is met by a line parallel to an axis in only a finite number of points and segments. If this condition is not satisfied, it is easy to show that in every neighborhood of P a point can be found in whose neighborhood E is met by a line parallel to an axis in only a finite number of points and segments. For the sake of complete symmetry, the equations defining E will be taken in parametric form

$$x = g(t), \quad (t \text{ in } T), \quad \sum A_\alpha^2 \neq 0,$$

the functions g being of class C' and A_α being given by (2:2).

Since the functions $y = y(x)$ minimize (2:1) we must have

$$I''(0) = \int_S 2\Omega(x, \eta, q)dx \geq 0. \tag{4:1}$$

This fact suggests a new problem of the calculus of variations, namely that of minimizing the integral (4:1) in the class of admissible manifolds $\eta = \eta(x)$. But it is clear from (3:1) that the functions

$$\begin{aligned} \eta &= u \text{ on } \bar{S} + \bar{M} \\ &= 0 \text{ elsewhere} \end{aligned} \tag{4:2}$$

give to $I''(0)$ its minimum value zero. Moreover, these functions

constitute a solution of class D" for the problem of minimizing I"(0) and hence must satisfy near P the edge conditions

$$\Omega_k^\alpha A_\alpha \Big|^1 = \Omega_k^\alpha A_\alpha \Big|^2 \qquad (k = 1, \cdots, m).$$

It is evident from (4:2) that the right members of these equations vanish, leaving

(4:3) $$\Omega_k^\alpha (x, u, w) A_\alpha \Big|^1 = 0 \text{ on } E, \qquad (k = 1, \cdots, m),$$

where $w_k^\alpha = (\partial u_k / \partial x_\alpha)$. In addition to (4:3) we have at P the equations

(4:4) $$du_1 = w_1^\alpha dx_\alpha = w_1^\alpha g_\alpha^\gamma dt_\gamma = 0.$$

Since the differentials dt_γ are arbitrary, (4:4) yields the equations

$$w_1^\alpha g_\alpha^\gamma = 0 \qquad (\gamma = 1, \cdots, n - 1).$$

From this we conclude that $w_1^\alpha = \lambda_i A_\alpha$; and (4:3) gives at P

$$\lambda_i \Omega_i^\alpha A_\alpha = \Omega_i^\alpha w_i^\alpha = f_{ik}^{\alpha\beta} \lambda_i A_\alpha \lambda_k A_\beta = 0.$$

But this contradicts the first hypothesis because the n vectors $w_1^\alpha, \cdots, w_m^\alpha$ constitute a non-zero set of collinear vectors. Thus we conclude that the admissible set of functions (4:2) does not give I"(0) its minimum value, and that, therefore I"(0) can be made negative, which, of course, precludes a minimum for I.

PART II

THE ANALOGUE OF JACOBI'S CONDITION FOR THE PARAMETRIC PROBLEM

5. Homogeneity conditions and edge conditions. The notations and definitions of Section 1 will be used in this part of our paper, with the exception that the independent variables x_α will be replaced here by the independent parametric variables u_α. It is assumed that the number m of variables y_i is greater than the number n of parameters u_α.

We consider the integral

$$I = \int_S F(y,\ p)du, \tag{5:1}$$

where F is of class C''' in W. A necessary and sufficient condition for the invariance of this integral under change of parameters is that F be expressible as a function $\Phi(y, J)$ of the y's and of the Jacobians J of the y's with respect to the u's, Φ being positively homogeneous of degree one in the J's.[1] Consequences of this condition are the identities

$$\begin{aligned} F_i^\alpha p_i^\beta &= F \quad \text{for } \alpha = \beta\ , \\ &= 0 \quad \text{"}\ \ \alpha \neq \beta\ . \end{aligned}$$

From these follow the further identities

$$\begin{aligned} F_{ik}^{\alpha\beta} p_i^\gamma &= 0 \quad \text{for } \alpha = \gamma\ ,\ \beta = \gamma\ , \\ &= 0 \quad \text{"}\ \ \alpha \neq \gamma\ ,\ \beta \neq \gamma\ , \end{aligned} \tag{5:2}$$

[1] Grosz, Monatshefte fur Mathematik und Physik, vol. 27 (1916), p. 70.

$$(5:2)\qquad \begin{aligned} F^{\alpha\beta}_{1k}\, p^{\gamma}_{1} &= F^{\beta}_{k} \quad \text{for } \alpha = \gamma\ ,\ \beta \neq \gamma\ , \\ &= -\, F^{\alpha}_{k} \quad \text{"} \quad \alpha \neq \gamma\ ,\ \beta = \gamma\ . \end{aligned}$$

In the following paragraphs we are to consider the problem of finding among all admissible manifolds $y = y(u)$, u in S, which pass through L, one which minimizes the n-tuple integral (5:1). The edge conditions for this problem are as stated in Theorem II, u replacing x.

Proceeding exactly as in Section 3, we obtain, if the functions $y = y(u)$ are to minimize (5:1),

$$I'(0) = 0$$

$$I''(0) = \int_S 2\Omega(u, \eta, q)du = \int_S \eta_i\left(\Omega_{\eta_i} - \frac{\partial \Omega^{\alpha}_{i}}{\partial u_{\alpha}}\right)du \geq 0,$$

2Ω being a quadratic form in the arguments η, $q^{\alpha}_{i} \equiv \dfrac{\partial \eta_i}{\partial u_{\alpha}}$. Here the Jacobi equations are the m equations

$$(5:3)\qquad \Omega_{\eta_i} - \frac{\partial \Omega^{\alpha}_{i}}{\partial u_{\alpha}} = 0\,, \qquad (i = 1, \cdots, m).$$

This system is of the second order and linear and homogeneous in the η's and their first and second derivatives. However, by virtue of the identities (5:2), the equations are not independent.

6. The Legendre quadratic form. Before proceeding to the proof of Jacobi's condition, two lemmas are proved for the purpose of throwing more light upon the first hypothesis of Theorem IV of the following section, which may be regarded as an analogue of Legendre's condition in the stronger form.

Denote by $V(p)$ the linear vector space determined by the n m-dimensional vectors

$$p^{\alpha}_{1} \cdots p^{\alpha}_{m}\,, \quad \alpha = 1, \cdots, n.$$

Each point of the manifold $y = y(u)$ has associated with it such

a space, namely the tangent manifold at that point.

LEMMA II. The quadratic form

$$F^{\alpha\beta}_{ik} h^{\,i}_{\alpha} h^{\,k}_{\beta} \tag{6:1}$$

vanishes for every set of n collinear vectors $(h^{\,1}_{\alpha}, \dots, h^{\,m}_{\alpha})$, $(\alpha = 1, \cdots, n)$ in $V(p)$.

The form, of course, vanishes for a set of zero vectors. Let $a_i = c_\alpha p^\alpha_i$ be any vector a in $V(p)$. Then $\bar{c}_\beta a_i$ represents the i-th component of the β -th vector of an arbitrary set of vectors collinear with a, the sets c and $\bar{c}$ being non-null. Let, further, i_1, α_1, β_1, be any fixed values of i, α, β in the respective ranges of the latter. We show that the expression

$$F^{\alpha_1\beta_1}_{i_1 k} h^{\,i_1}_{\alpha_1} h^{\,k}_{\beta_1} + F^{\beta_1\alpha_1}_{i_1 k} h^{\,i_1}_{\beta_1} h^{\,k}_{\alpha_1} \tag{6:2}$$

vanishes for $h^{\,i}_{\alpha} = \bar{c}_\alpha a_i$, and that therefore (6:1) also vanishes. Under the substitution $h^{\,i}_{\beta} = \bar{c}_\beta a_i = \bar{c}_\beta c_\alpha p^\alpha_i$, (6:2) becomes

$$h^{\,i_1}_{\alpha_1} \bar{c}_{\beta_1} F^{\alpha_1\beta_1}_{i_1 k} c_\alpha p^\alpha_k + h^{\,i_1}_{\beta_1} \bar{c}_{\alpha_1} F^{\beta_1\alpha_1}_{i_1 k} c_\alpha p^\alpha_k . \tag{6:3}$$

Since $h^{\,i_1}_{\alpha_1} \bar{c}_{\beta_1} = h^{\,i_1}_{\beta_1} \bar{c}_{\alpha_1} = K$, we see by reference to the identities (5:2) that the expression (6:3) vanishes for $\alpha_1 = \beta_1$ and becomes

$$K(c_{\beta_1} F^{\alpha_1}_{i_1} - c_{\alpha_1} F^{\beta_1}_{i_1} + c_{\alpha_1} F^{\beta_1}_{i_1} - c_{\beta_1} F^{\alpha_1}_{i_1}) = 0$$

for $\alpha_1 \neq \beta_1$

LEMMA III. If the quadratic form

$$F^{\alpha\beta}_{ik} h^{\,i}_{\alpha} h^{\,k}_{\beta} \tag{6:4}$$

vanishes for a set of n collinear vectors not in $V(p)$, then it vanishes for a non-zero set of collinear vectors orthogonal to $V(p)$.

Let $\bar{c}_\beta b_i$ represent the i-th component of the β-th vector of a set of collinear vectors not in V(p) for which (6:4) vanishes. There exists a vector a in V(p) and a vector $\bar{a}$ orthogonal to V(p) such that

$$\bar{a} = b + a.$$

Then the form vanishes for the set

$$\bar{c}_\beta \bar{a} = \bar{c}_\beta (b + a).$$

Place $h^i_\alpha = \bar{c}_\alpha \bar{a}_i = \bar{c}_\alpha b_i + \bar{c}_\alpha a_i$. Then (6:4) becomes

$$\begin{aligned} (6{:}5) \qquad & F^{\alpha\beta}_{ik} (\bar{c}_\alpha b_i + \bar{c}_\alpha a_i)(\bar{c}_\beta b_k + \bar{c}_\beta a_k) \\ &= F^{\alpha\beta}_{ik} \bar{c}_\alpha \bar{c}_\beta b_i b_k + F^{\alpha\beta}_{ik} \bar{c}_\alpha \bar{c}_\beta b_i a_k + F^{\alpha\beta}_{ik} \bar{c}_\alpha \bar{c}_\beta a_i b_k + F^{\alpha\beta}_{ik} \bar{c}_\alpha \bar{c}_\beta a_i a_k. \end{aligned}$$

The first and fourth sums vanish by hypothesis and Lemma II. Replacing in the second sum a_k by $c_\alpha p^\alpha_k$ and setting $i, \alpha, \beta = i_1, \alpha_1, \beta_1$ we form expressions analogous to (6:2) and (6:3) and thus show that this also vanishes. But the second and third sums are clearly equal. Therefore the expression (6:5) vanishes.

7. The analogue of Jacobi's condition.

Definition VIII. A solution v of the Jacobi equations (5:3) is called a normal solution if it satisfies identically the relations

$$(7{:}1) \qquad v_i p^\alpha_i = 0$$

on the manifold $\mathfrak{M}$.

We shall have occasion to use the derivatives of (7:1),

$$(7{:}2) \qquad \pi^\beta_i \, p^\alpha_i + \text{linear terms in } v_i = 0$$

where $\pi^\beta_i = \partial v_i / \partial u_\beta$.

THEOREM IV. Jacobi's Condition.

Hypotheses:

1. $\mathfrak{M}$ is a minimizing manifold of class C" on which the Legendre quadratic form

$$F^{\alpha\beta}_{ik} h^i_{\alpha} h^k_{\beta}$$

is different from zero for all non-zero sets of collinear vectors orthogonal to V(p).

2. $\overline{M}$ is any (n - 1)-dimensional Green's manifold bounding a domain $\overline{S}$ such that $\overline{S}$ is in the domain of integration S of U-space and the point set S - ($\overline{S}$ + $\overline{M}$) consists of a finite number of sets of type S.

3. There exists a normal solution v of (5:3) which vanishes on $\overline{M}$, is of class C" on $\overline{S}$ + $\overline{M}$, and is such that the partial derivatives π are not all zero at a point P of $\overline{M}$ in S.

Conclusion:

The manifold $\mathfrak{M}$ does not yield a minimum for the integral I.

As in Section 4 it can be shown easily that the portion E of $\overline{M}$ in the neighborhood of P can be represented by equations of the form

$$u = g(t), \quad (t \text{ in } T), \quad \sum A^2_{\alpha} \neq 0\,,$$

where the g are of class C' and A_{α} is given by (2:2). Moreover, without restricting ourselves we may assume that in a neighborhood of P, E is met by a line parallel to an axis in at most a finite number of points and segments. We argue, as before, that if y = y(u) furnish a minimum, we must have $I''(0) \geq 0$, and that, since the admissible functions

$$\eta = v \text{ on } \bar{S} + \bar{M} \qquad (7:3)$$
$$= 0 \text{ elsewhere}$$

give to $I''(0)$ its minimum value zero, they must constitute a solution of class D'' for the problem of minimizing the integral

$$J = I''(0) = \int_S 2\Omega(u, \eta, q)du. \qquad (7:4)$$

Hence, the edge conditions

$$\Omega^{\alpha}_{k}(u, v, \pi)A_{\alpha}\Big|^{1} = 0 \text{ on } E, \quad (k = 1, \cdots, m) \qquad (7:5)$$

must be satisfied. Also along E we have

$$dv_i = \pi^{\alpha}_{i}\, du_{\alpha} = \pi^{\alpha}_{i}\, g^{\gamma}_{\alpha} dt_{\gamma} = 0,$$

which imply the equations

$$\pi^{\alpha}_{i}\, g^{\gamma}_{\alpha} = 0, \qquad (\gamma = 1, \cdots, n - 1). \qquad (7:6)$$

From the latter it follows that

$$\pi^{\alpha}_{i} = \lambda_i A_{\alpha}. \qquad (7:7)$$

Since $v = 0$ on E, (7:5) yields at P the equation

$$\Omega^{\alpha}_{k}(u, v, \pi)\,\pi^{\alpha}_{k} = 0 = F^{\alpha\beta}_{ik}\,\pi^{\alpha}_{i}\,\pi^{\beta}_{k}.$$

Concerning the n m-dimensional vectors π^{α}_{i}, we observe that at P they are not all zero by hypothesis 3; they are collinear by equations (7:7); and they are orthogonal to V(p) by (7:2). We conclude, therefore, that the functions (7:3), which cause the second variation to vanish do not minimize it, and that consequently there exists an admissible set of variations for which $I''(0) < 0$.

PART III

A FURTHER RESULT

For the simple integral

$$I = \int_{x_1}^{x_2} f(x, y, z, y', z')dx$$

Jacobi's condition may be stated as follows: If E_{12} is a minimizing arc having no corners and such that along it $f_{y'y'}f_{z'z'} - (f_{y'z'})^2$ is different from zero, and if x_3 is a point between x_1 and x_2, then there can exist no solution of Jacobi's equations which vanishes at x_1 and at x_2 and which is not identically zero between x_1 and x_2. For the double integral

$$I = \iint_S f(x, y, z, p, q)dx\, dy,$$

the following analogous statement is found in the literature: If Legendre's condition in the stronger form holds along a minimizing surface of class C", then there can exist no solution u of Jacobi's equation of class C" which vanishes along a Green's curve $\overline{M}$ in S, and which is not identically zero in the domain $\overline{S}$ bounded by $\overline{M}$. These statements are somewhat more illuminating than our Theorem III because they do not postulate the non-vanishing of a derivative at the conjugate point x_3 or at a point of the curve $\overline{M}$. In the case of the simple integral this elegance of form of statement is readily accounted for by the fact that if a solution together with its derivatives did vanish at x_3, it would vanish identically on the interval (x_1, x_3) by virtue of the Cauchy existence theorem for a system of non-analytic ordinary differential equations. But the method used to obtain the

corresponding result for the double integral involves an altogether different set of ideas and has not yet been sufficiently developed to yield an analogous result for the problems considered in this paper.

We recall briefly the latter method for a well-known case[1]. To the equation

$$(1) \qquad L(u) \equiv \partial^2 u/\partial x^2 + \partial^2 u/\partial v^2 + a(\partial u/\partial x) + b(\partial u/\partial y) + cu = q$$

adjoin the equation

$$M(v) \equiv \partial^2 v/\partial x^2 + \partial^2 v/\partial y^2 - \partial(av)/\partial x - \partial(bv)/\partial y + cv = 0.$$

Then it follows readily that

$$\iint_{\bar{S}} \left[vL(u) - uM(v) \right] dx\, dy = - \int_{\bar{M}} \Big\{ v(\partial u/\partial n) - u(\partial v/\partial n) + (a \cos(nx) + b \cos(ny))uv \Big\} ds.$$

If a, b, c are of class C" in $\bar{S} + \bar{M}$, then corresponding to a point (ξ, η) of $\bar{S}$ the adjoint equation admits an elementary solution of the form

$$v = U(x, y; \xi, \eta) \log r + V(x, y; \xi, \eta), \qquad r = \sqrt{(x - \xi)^2 + (y - \eta)^2},$$

U and V being of class C" and $U(\xi, \eta; \xi, \eta) = -1$. A consequence of the existence of this elementary solution is the important formula for the value of a solution u of equation (1) at a point of $\bar{S}$,

---

[1] Picard, Equations aux derivees partielles, p. 192 ff.
Webster-Szego, Partielle Differentialgleichungen, p. 292.
Goursat, Cours d'Analyse III, pp. 230 ff.

$$(2) \quad 2\pi u(\xi,\eta) = -\int_M \left\{ v\frac{\partial u}{\partial n} - u\frac{\partial v}{\partial n} + (a\cos(nx) + b\cos(ny))uv \right\} ds - \iint_{\overline{S}} vg\, dx\, dy,$$

where $\partial u/\partial n$ and $\partial v/\partial n$ are taken along the interior normal of M and cos (nx) and cos (ny) are the direction cosines of the latter.

It is seen that the method consists of three steps: (1) an approoriate definition of the adjoint in order to secure a generalization of Green's formula; (2) the proof of the existence of an elementary solution; (3) the derivation by means of the elementary solution of a formula which yields the value of a solution at a point of $\overline{S}$ in terms of its values and those of its partial derivatives (first partial derivatives for an equation of the second order) on the boundary M. An important fourth step is the determination of Green's function; that is, an elementary solution that vanishes on the boundary, or satisfies some other boundary condition suggested by the problem in hand. This does not concern us here.

Consider the general second order equation

$$(3) \quad F(u) \equiv \sum_{\alpha,\beta=1}^{n} a_{\alpha\beta}\frac{\partial^2 u}{\partial x_\alpha \partial x_\beta} + \sum_{\alpha=1}^{n} b_\alpha \frac{\partial u}{\partial x_\alpha} + cu = f, \quad a_{\alpha\beta} = a_{\beta\alpha},$$

and its adjoint

$$G(v) \equiv \sum_{\alpha,\beta=1}^{n} \frac{\partial^2 (a_{\alpha\beta} v)}{\partial x_\alpha \partial x_\beta} - \sum_{\alpha=1}^{n} \frac{\partial (b_\alpha v)}{\partial x_\alpha} + cv = 0$$

From the identities

$$v a_{\alpha\beta}\frac{\partial^2 u}{\partial x_\alpha \partial x_\beta} - u\frac{\partial^2 (a_{\alpha\beta} v)}{\partial x_\alpha \partial x_\beta} = \frac{\partial}{\partial x_\alpha}\left(v a_{\alpha\beta}\frac{\partial u}{\partial x_\beta}\right) - \frac{\partial}{\partial x_\beta}\left(u\frac{\partial (v a_{\alpha\beta})}{\partial x_\alpha}\right),$$

$$v b_\alpha \frac{\partial u}{\partial x_\alpha} + u\frac{\partial (v b_\alpha)}{\partial x_\alpha} = \frac{\partial (v u b_\alpha)}{\partial x_\alpha}$$

the further identity

$$vF(u) - uG(v) = \sum_1^n \frac{\partial P_\alpha}{\partial x_\alpha}$$

follows, where

$$P_\alpha = v\sum_\beta a_{\alpha\beta}\frac{\partial u}{\partial x_\beta} - u\sum_\beta \frac{\partial}{\partial x_\beta}(v\, a_{\alpha\beta}) + u v b_\alpha \,.$$

Setting

$$\pi_\alpha = \cos(n, x_\alpha), \quad \frac{\partial u}{\partial T} = \sum_{\alpha,\beta} a_{\alpha\beta}\pi_\alpha \frac{\partial u}{\partial x_\beta}\,; L = \sum_\alpha \pi_\alpha\Big(b_\alpha - \sum_\beta \frac{\partial a_{\alpha\beta}}{\partial x_\beta}\Big),$$

we find, by Green's theorem,

$$\int_{\bar S}\big[v F(u) - u G(v)\big]d\bar S = -\int_{\bar M}\Big[v\frac{\partial u}{\partial T} - u\frac{\partial v}{\partial T} + L u v\Big]d\bar M.$$

The equation (3) is elliptic in $\bar S + \bar M$ if at each point of the region the form $\sum a_{\alpha\beta}\sigma_\alpha\sigma_\beta$ is positive definite. For convenience take $|a_{\alpha\beta}| \equiv 1$. The existence of an elementary solution of this equation was proved by Hadamard (18) for the analytic case, and by Levi (19) under the assumption that the $a_{\alpha\beta}$ are of class C''' and the b_α, c, f of class C'. Levi reduces the problem to the solution of an integral equation of Fredholm type. But his discussion at this point is quite brief. Some of the desired details are to be found in Sternberg (22) and Gevrey (32). For $n > 2$ an elementary solution is given under the form

$$v(P,\pi) = \Big[\sum_1^n a^{\alpha\beta}(P)(x_\alpha - \xi_\alpha)(x_\beta - \xi_\beta)\Big]^{-\frac{n}{2}+1} + W(P,\pi),$$

where $\sum a^{\alpha\beta}\sigma_\alpha\sigma_\beta$ is the form adjoint to $\sum a_{\alpha\beta}\sigma_\alpha\sigma_\beta$ and $W(P,\pi)$ is such that in the neighborhood of the point π,

$$|W(P,\pi)| < \mathrm{const}/|P\pi|^{n-3}, \quad |\partial W(P,\pi)/\partial x_\alpha| < \mathrm{const}/|P\pi|^{n-2};$$

while for $n = 2$ a logarithmic expression forms the principal part. The existence of such solutions for the adjoint equation makes possible the derivation (32, p. 9) of the fundamental formula for the value of a solution u of equation (3),

$$
(4)\qquad \begin{aligned} s(n-2)u(\Pi) = -\int_{\overline{M}} \left[v(\partial u/\partial T) - u(\partial v/\partial T) + Luv \right] d\overline{M} - \\ \int_{\overline{S}} vf \, d\overline{S}, \end{aligned}
$$

s being the surface of the unit hypersphere.

For the problem of minimizing the integral

$$
(5)\qquad I = \int^{n} f(x_1, \cdots, x_n, y, p_1, \cdots, p_n)dx_1 \cdots dx_n,
$$

where $p_\alpha = \partial y/\partial x_\alpha$, we have the single Jacobi equation

$$
\Omega_y - \sum_{\alpha=1}^{n} \frac{\partial \Omega}{\partial x_\alpha} q_\alpha = 0, \; q_\alpha \equiv \frac{\partial u}{\partial x_\alpha}.
$$

This equation is linear and homogeneous in u and its first and second partial derivatives and is of the form $F(u) = 0$. Moreover it is self-adjoint and in view of Legendre's condition in the stronger form it is elliptic in $\overline{S} + \overline{M}$. Hence, under such conditions as validate formula (4) we may for the problem (5) announce Jacobi's condition in the form suggested at the beginning of this section.

The extension of the method outlined above to elliptic systems of equations offers difficulties with respect to each of the three steps, particularly the second step. Levi used his method to prove the existence of an elementary solution of a non-analytic elliptic system in two independent variables, while Cioranesco (33) gives methods of adjunction for the system

$$
L_\alpha \equiv \operatorname{div}(\operatorname{grad} u_\alpha) + \sum_{\beta=1}^{n}(a_{\alpha\beta} \operatorname{grad} u_\beta + b_{\alpha\beta} u_\beta) = f_\alpha, \; (\alpha = 1, \cdots, n),
$$

and, assuming the existence of elementary solutions, derives formulas analogous to (4). The extension of this important method to the general elliptic system with an arbitrary number of independent variables is still an open problem.

BIBLIOGRAPHY

LEGENDRE'S CONDITION FOR MULTIPLE INTEGRALS

1. Hadamard, J., "Sur une question de calcul des variations," Bull. de la Soc. Math. de France, vol. 30 (1902), p. 253.

2. Hadamard, J., "Sur quelques questions de calcul des variations," Bull. de la Soc. Math. de France, vol. 33 (1905), p. 73.

3. Caratheodory, C., "Ueber die Variationsrechnung bei Mehrfachen Integralen," Acta Litt. ac Scient., vol 4 (1929), p. 193.

4. McShane, E., "Weierstrass Condition in the Multiple Integral Problem," Annals of Math., vol. 32 (1931), p. 578.

JACOBI'S CONDITION FOR DOUBLE INTEGRALS

5. Kobb, G., "Sur les maximas et les minimas des integrales double," Acta Math., vol. 17 (1892), p. 65.

6. Sommerfeld, A., "Bemerkungen zur Variationsrechnung," Deutsche Math. Verein, vol. 8 (1899), p. 188.

7. Kurosu, K., "Some Fundamental Theorems in the Calculus of Variations," Tohoku Math. Journal, vol. 7 (1915), p. 54.

8. Kubota, T., "Zur Variationsrechnung," Tohoku Sci. Reports, vol. 5 (1916), p. 71.

9. Kubota, T., "Die Jacobische Bedingung in der Variationsrechnung," Tohoku Science Reports, vol. 5 (1916), p. 241.

10. Kubota, T., "Zur Variationsrechnung," Tohoku Math. Journal, vol. 10 (1916), p. 46.

11. Lichtenstein, "Uber zwei-dimensionalen regulare Variationsprobleme," Monat. fur Math. und Physik, vol. 28 (1917), p. 3.

12. Lichtenstein, L., "Untersuchungen uber zwei-dimensionale Variationsprobleme," Math. Zeit., Bd. 5, (1919), p 26.

13. Picone, M., "Nuova condizione necessaria per un estrema di un integrale doppio," Atti dei Lincei, vol. 5 (1922), p. 94.

14. Simmons, H., "First and second variations of a double integral for the case of variable limits," Trans., vol. 28 (1926), 235.

15. Haar, A., "Zur Variationsrechnung," Abh. aus dem Math. Seminar der Hamburgische Universitat, vol. 8 (1930).

ELLIPTIC PARTIAL DIFFERENTIAL EQUATIONS WITH THREE OR MORE INDEPENDENT VARIABLES

16. Cotton, E., "Invariants differentiel de quelques équations lineares," Ann. Sc. de l'Ec., 17 (1900), p. 211.

17. Holmgren, E., "Existenz der Grundlösung," Math. Annalen, vol. 58 (1904), p. 404.

18. Hadamard, J., "Sur les solutions fondamentales des équations linéares," Ann. Sc. de l'Ec., Norm. Sup., (1904), p. 535; (1905), p. 101.

19. Levi, E.E., "Equazioni lineari totalmente ellittiche," Rend. del. circ. mat. di Palermo, (1907), p. 275.

20. Fubini, G., "Soluzioni fondamentali," Rend. Ac. Lincei, vol. 18 (1909), p. 423.

21. Hadamard, J., Lectures on Cauchy's Problem, Yale U. Press,1923.

22. Sternberg, W., "Lineare, elliptische Differentialgleichungen," Math. Zeit., vol. 21 (1924), p. 286.

23. Lichtenstein, L., "Neue Beitrage zur Theorie der linearen Differentialgleichungen," Math. Zeit., vol. 20 (1924).

24. Giraud,G., "Problème di Dirichlet généralise," Annales Ec. norm., vol. 43 (1926), p. 1.

25. Picone, M., "Maggiorazione degli integrali delle equazioni lineari," Atti. d. r. Ac. dei Lincei, vol. 5, fasc. 3, 1927.

26. de Donder, J., "Sur l'équation d'un ordre quelconque," Journal d. math. p. et appl. (1928), p. 173.

27. Rothe, E., "Eigenfunktionen elliptischer Differentialgleichungen," Sitzungsberischte B.M.G., (1929), p. 71.

28. Cioranescu, N., "Le Problème de Dirichlet pour les systèmes d'équations," Theses, Paris, 1929.

29. Giraud, G., "Équations de type elliptique et la methode des approximations successive," Jour. de Math., (1929), p. 269.

30. Giraud, G., "Le Problème de Dirichlet généralise," Ann. Ecole. norm., (1929), p . 131.

31. Giraud, G., "Les équations du type elliptique," Bull. sc. Math., (1929), p. 367.

32. Gevrey, M., "Détermination et emploi des fonctions de Green, Journal de Math. p. et appl., (1930), p. 1.

33. Cioranescu, N., "Le Problème de Dirichlet," Mathematica, (1930), p. 67.

34. Feller, W., "Lösungen der linearen Differentialgleichungen," Math. Annalen, vol. 104 (1931), p. 636.

A HISTORY OF THE CLASSICAL ISOPERIMETRIC PROBLEM

BY

THOMAS ISAAC PORTER

TABLE OF CONTENTS

A HISTORY OF THE CLASSICAL ISOPERIMETRIC PROBLEM

<u>Introduction</u>. The famous old Isoperimetric Problem of the ancients was that of finding a simply closed curve of given length which incloses the largest area. Another problem closely related to this problem is that of finding among all curves which inclose a given area that one which has the shortest perimeter. It is easy to prove that the solution to either of the two problems leads logically to the solution of the other. The first problem may be stated analytically as that of finding an arc with equations in parametric form

$$x = x(t) \ , \quad y = y(t) \ , \quad t_1 \leq t \leq t_2$$

which satisfies the conditions

$$x(t_1) = x(t_2) \ , \quad y(t_1) = y(t_3) \ ,$$

but does not otherwise intersect itself, which gives the length integral

$$\int_{t_1}^{t_2} \sqrt{x'^2 + y'^2} \, dt$$

a fixed value l, and maximizes the area integral

$$1/2 \int_{t_1}^{t_2} (xy' - yx')\,dt.$$

The solution is a circle. The general problem of the calculus of variations for which one integral is to be given a fixed value while another is to be made a maximum or minimum is called after this one an isoperimetric problem. Any problem where a fixed length is involved while an integral of any kind is to be made a maximum or minimum is called an isoperimetric problem. In this paper only the first two problems which are formulated above will be discussed.

The earliest attempted solution of the problem which has been preserved for us was that of Zenodorus. His solution was the one generally given for the problem until the time of Steiner. Steiner gave a very elegant and simple proof of a condition necessary for a solution but did not give a sufficiency proof. The first complete proof that the solution of the problem is a circle was given by Weierstrass. The earliest writers who attempted to solve the problem used a geometric method. Later complete theories have been given by the method of the calculus of variations, by the geometric method, and more recently by means of Fourier Series.

The discussion of the history of the Isoperimetric Problem is divided below into four sections. In Section I the origin of the problem and the proofs to the beginning of

the nineteenth century are discussed. In Section 2 the development of the geometric proofs is traced from the beginning of the nineteenth century to the present. The greater part of Section 2 is concerned with the development of the theory of Steiner, and of the theory of parallel curves which was first applied to the solution of the Isoperimetric Problem by Minkowski. Section 3 consists of an outline of the proofs of the Isoperimetric Property of a circle which have been given by means of Fourier Series. In Section 4 solutions which have been given by means of the calculus of variations are discussed. The greater part of Section 4 is a discussion of the application of the theory of Weierstrass to the Isoperimetric Problem.

The Bibliography which follows Section 4 is also divided into four groups to correspond to the four sections described in the preceding paragraph. References which merely mention the Isoperimetric Problem are not included.

The numbers in the text inclosed in parentheses and following an author's name, refer to the dates of birth and death of the author. The numbers inclosed in square brackets refer to the Bibliography.

1. <u>Origin of the problem and early theories</u>. The origin of the problem has been attributed to the Greeks but it is not known who among them was the first to state the problem or attempt a solution. The mathematical historians Montucla [4] and Cantor [3] quote a statement of Diogenes Laertius (third century) regarding Pythagoras (580?-500? B.C.) which has been interpreted to mean that Pythagoras knew the maximum property of the circle. The statement of Pythagorus is, however, not at all convincing as it merely asserts that of all plane figures the circle is the most beautiful.

The following remark [9] appears in the De Caelo of Aristotle (384-322 B.C.): "Now of lines which return upon themselves the line which bounds the circle is the shortest." Aristotle does not give any further explanation of the sentence and his subject is a philosophical rather than a mathematical one.

Archimedes (287-212 B.C.) may have been the first to attempt a mathematical solution of the problem. Simplicius who lived about the sixth century A.D. mentions a proof [14] of Archimedes and Zenodorus. However, Proclus (410-485) says [12] that later mathematicians arrived at a solution partly from the works of Euclid (450-374 B.C.) and partly from those of Archimedes. The mathematical historian Libri [6] gives a list of books found in the Cosmography of Maurolytus (1494-1575) and one called the isoperimetric

Figures appears under the name of Archimedes. Isaac Barrow (1630-1677) in his preface [20] to the Book of Lemmas makes the following comment in a note on the margin of his book: "These men attribute to him (Archimedes) a book which is the Isoperimetry of Zenodorus, fragments of statics, and other works."

It is not clear whether the references of the last paragraph are to a work of Archimedes which has been lost or to the later collection of articles called the Book of Lemmas to which the name of Archimedes has by some critics thought to have been erroneously appended. Zenodorus in one of his proofs [16 Hultsch ed., vol. 3, p. 1194] used a theorem which he says was a theorem of Archimedes. This is evidence for a conclusion that Zenodorus was acquainted with the work of Archimedes, and that he would have mentioned the fact if Archimedes had arrived at a solution of the Isoperimetric Problem.

A statement [10] of the historian Polybius (201-120 B.C.) suggests that he may have known the maximum property of the circle. He says that most people judge the size of cities simply from their circumference and that when one tells them that a city or camp with a circumference of forty stades may be twice as large as one with a circumference of one hundred stades the statement seems astounding to them. He adds that the trouble is that we have forgotten our lessons in geometry.

The Greeks in general did not have a clear understanding regarding the relation of perimeter to area. Polybius adds to the above remark that not only ordinary men but also statesmen and military men judged the size of a camp by measuring the circumference. Proclus mentions [13] certain members of communistic societies who cheated their fellow members by giving them plots of land of larger perimeter but smaller area than the plots which they took for themselves. He says also that the theorem that all triangles formed on the same or equal bases, and always between the same two parallel lines are equal in area was regarded by the Greeks as paradoxical because the perimeter could be made as large as one pleased.

Zenodorus who lived probably during the period from 200 B.C. to 90 A.D. wrote a treatise on figures of equal perimeter. This work has been lost but has been partly preserved in the works [15,16] of Theon of Alexandria and Pappus who were contemporaries and probably lived during the fourth century A.D. Christopher Clavius (1537-1612) refers [17] to the works of Theon and Pappus as the sources from which he took the proofs which he gives regarding isoperimetric figures. Recently discussions of the theorems of Zenodorus have been given by Heath [7] and Chisini [8] but both works follow that of Pappus. The theorems of Zenodorus that are of interest for our problem are:

Theorem I. Among all polygons of equal number of sides and equal perimeters, the regular polygon is greatest in area.

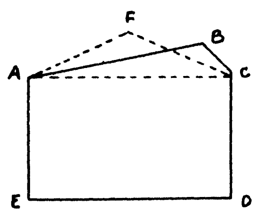
Theorem II. The circle is greater in area than any regular polygon which has an equal perimeter.

The proof of Theorem I depends on two lemmas.

Lemma I. Among all triangles having the same base and the same sum of sides, the isosceles triangle is greatest in area.

Lemma II. When two isosceles triangles are not similar to each other, if we construct on the same bases two triangles that are similar to each other, and such that the sum of the perimeters of the similar triangles is equal to that of the two original triangles, then the sum of the areas of the similar triangles is greater than the sum of the areas of the non-similar triangles.

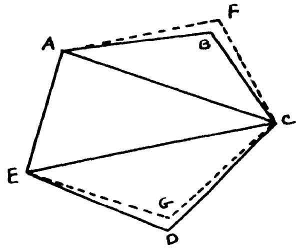
The proof of Theorem I according to Zenodorus is as follows: Suppose AB and BC are unequal. Construct AF equal to CF and such that the sum of AF and CF is equal to the sum of AB and BC. Then by Lemma I, the triangle AFC is greater than the triangle ABC. Therefore the area



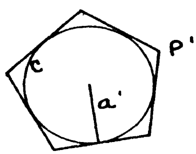
of the new figure is greater than the area of the original one, which is contrary to the hypothesis that the given figure

be a maximum. Therefore the maximum polygon must be equilateral.

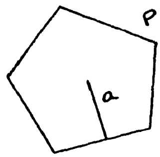
Now suppose the angle B is greater than the angle D, while the polygon is equilateral. Construct isosceles triangles AFC and EGC similar to each other and such that the sum of their perimeters is equal to the sum of the perimeters of ABC and EDC. Then by the use of Lemma II we know that the sum of the areas of the triangles AFC and EGC is greater than the sum of the areas of the triangles ABC and EDC. Therefore we have a new polygon of equal perimeter but greater area than the original one which is contrary to the hypothesis that the original polygon is a maximum polygon. By repeating the argument for other pairs of angles we finally conclude that the maximum polygon must be equiangular. Therefore the maximum polygon must be regular.



Zenodorus used the following demonstration as a proof of Theorem II. Let C be a circle of perimeter p, and P be a regular polygon of equal perimeter. Let P' be a polygon circumscribing C and similar to P. Let a and a' be the apothems of P and P', and



notice that a' is the radius of the circle. Since the polygons are similar we know that a/a' is equal to p/p'. But p' is greater than the perimeter p of the circle, therefore a' is greater than a. The area



of C is, by the use of a theorem proved by Archimedes, equal to one-half the area of a rectangle the length of which is the perimeter and the width the radius of C, or a'p/2, and the area of P is ap/2. Therefore the area of C is greater than the area of P.

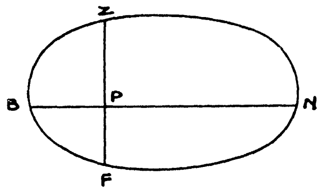
The two theorems when completely proved justify the conclusion that a circle of any given perimeter incloses a larger area than any polygon which has an equal perimeter. But in Theorem I it is assumed that among all polygons of a given perimeter there is one that is a maximum and this statement requires a proof. In addition we need for a complete proof of the Isoperimetric Theorem a discussion of the case when the figure which we are comparing with the circle is not a polygon.

Kepler (1571-1630) states the theorem [18] and refers the reader to the proof of Pappus. Galileo (1564-1642) gives an argument [19] which is essentially the proof of Zenodorus.

James Bernoulli (1654-1705) mentions the problem [21]

and says that the solution is a circle but adds the statement needs to be proved. On the same page he proposed to mathematicians a more general problem. His problem is as follows:

Among all curves BFN of equal length and having a common base BN, to find one such that for a corresponding curve BZN whose ordinate PZ is any function of the ordinate PF or of the arc BF, the area BZNB shall be a maximum.



If we choose PZ equal to PF then the solution is two equal arcs of a circle. The solution will be a complete circle if we further choose the sum of the lengths of the two curves to be π times the length of BN.

The numerous attempts to solve the problem of Bernoulli and others similar to it led to the development of a new method for solving such problems. This new method Euler called the calculus of variations. All the early writers on the calculus of variations were concerned only with proving the conditions necessary for a solution. As a result the early proofs of the maximum property of the circle by the new method as well as the older geometric method were not complete.

Euler (1707-1783) developed a theory which is called the rule of Euler and applied it [22] to proving that if a

curve of given length incloses a maximum area it is necessary that the curve be a circle. Riccati (1707-1775) has written a dissertation [26] in which he discusses the maximum property of the circle in a manner similar to that of Zenodorus which has been given in a preceding paragraph of this section. T. Simpson (1710-1761) gave a discussion [25] of the problem similar to that of Euler. He also gave a geometric demonstration [29] of the problem which is essentially the one given by Zenodorus. Other geometric proofs [23,30] much the same as that of Simpson have been given by Legendre (1752-1833) and P. Elvius (1710-1749). S. L'Huiler (1750-1840) proved [27] that if a curve which has a given length incloses a maximum area it is necessary that the radius of curvature be constant.

2. Geometric proofs of the isoperimetric property of the circle. The early attempts to prove the Isoperimetric Theorem by means of geometry lacked generality since only polygons were used as comparison curves. During the beginning of the nineteenth century an attempt was made to give a more general type of proof. The geometric proofs attempted from the beginning of the nineteenth century to the present can be roughly divided into two groups. The first group is largely synthetic and is closely associated with the work of Steiner (1796-1863). The second group is analytic and is concerned with proving an auxiliary theorem from which the Isoperimetric

Theorem follows at once. This auxiliary theorem says that for any closed curve the square of the perimeter minus 4π times the inclosed area is greater than or equal to zero and the equality sign is valid only when the curve is a circle.

An attempt to prove that the area of a circle is larger than that of any closed curve of equal perimeter appeared in 1813 in an article [32] by an unknown author. The writer assumes that among all plane figures inclosing equal areas there is one that has the smallest perimeter. He then proves that this figure of smallest perimeter must be a circle. He argues that if one denies that it is a circle then some other figure must have this property. In that case he shows that a new figure can be constructed having equal area but smaller perimeter than the original one. But such a construction contradicts the hypothesis that the original figure possesses the smallest perimeter possible. Therefore, among all curves of equal area the circle has the shortest perimeter. Now suppose that there is a figure S having an equal perimeter and a larger area than a given circle C'. Make a circle C equal in area to S. The perimeter of C will be less than the perimeter of S by the preceding theorem. The area of C' will be larger than the area of C and therefore larger than the area of S which is contrary to hypothesis.

Steiner gave five different proofs [33] of the maximum property of the circle. His theorems include the analo-

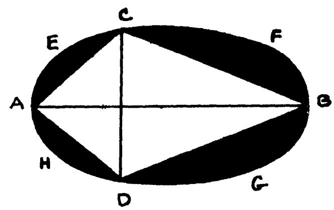
gous proposition that among all figures having equal areas a circle has the shortest perimeter. He states that each of these propositions implies the other as indicated for one of them in the last paragraph above, and therefore proves only one of them in any particular discussion. The proofs are general enough to include all closed curves having a given perimeter, but are in each case based on one of the following postulates which are not proved:

(1) Among all closed curves having a given perimeter there is at least one whose area is equal to or greater than the area of any of the others.

(2) Among all closed curves of given area there is at least one whose perimeter is less than or equal to the perimeter of any of the others.

Steiner comments on the first postulate as follows: "It is clear that there is an infinity of figures which have equal perimeters but which have different form and area. One observes that the area can be made as small as one pleases but not as large as one pleases, since all the figures may be inclosed in a circle whose radius is equal to one half the perimeter of the given figures and whose center is one point of the circumference of one of the figures. It must be that among these there is a maximum figure or several maximum figures, that is several figures which have equal perimeter but a larger area than any other figure not in the group."

__Theorem.__ __Among all curves of equal perimeter the circle incloses the maximum area.__



Let EFGH be a maximum figure. One can find a line AB that divides the perimeter into two equal parts. Therefore, the area is divided into two equal parts because if not one can replace the smaller half by one equal to the larger and thus increase the area of the original figure without changing the perimeter which is contrary to hypothesis. If the figure is not symmetric with respect to AB, replace one half of the figure by a figure symmetric to the other half, and since the area and perimeter are unchanged the figure is still one of the maximum figures. Now choose any point D on a semiperimeter. From it draw a perpendicular to AB and extend this perpendicular to meet the perimeter again at C, and draw the quadrilateral ACBD. The angles of the quadrilateral at C and D are right angles, because if not one can transform the figure so as to make them right angles while keeping the parts between the quadrilateral ACBD and the perimeter fixed, thus increasing the area without changing the perimeter which is contrary to hypothesis. Since D is an arbitrary point the figure must be a circle.

Since the other proofs are quite similar to the above, only a brief summary of the two proofs most discussed by writers following Steiner will be given. These are his second and fifth proofs. The second method introduces a quadrilateral drawn through any four points of the perimeter and the use of the fact that the maximum polygon of given perimeter is one that is inscribed in a circle in order to prove that for any figure other than a circle the area can be increased when the perimeter is kept fixed. The fifth method shows that a maximum figure must be symmetric to an arbitrary axis and therefore must be a circle.

Edler attempted to make a proof [34] by describing a geometric construction for making a regular polygon of at most 2^{n-1} sides from an irregular polygon of n sides. The constructed polygon has a smaller perimeter than the given polygon and a surface area at least equal to the surface area of the given polygon. He also proves that a circle has a larger surface area than any regular polygon which has an equal perimeter. These two theorems are combined into a proof of the maximum property of the circle as compared to polygons. To take care of the case when the figure is not a polygon the method of Steiner is used to increase the area while the perimeter remains constant. A polygon whose perimeter is equal to the perimeter of the original figure and which has also an equal area is inscribed in the increased

figure. The theory regarding polygons is then applied to complete the proof.

Sturm (1841-1919) gave a discussion [35] of the maximum property of the circle. His proof of the condition necessary for a maximum is the same as that of Steiner. He has also simplified the construction of Edler described above.

Witting (1861-) proved [39] that any figure, such that if a line divides its perimeter into two equal parts, its area is also divided into two equal paets, is a figure with a center point. He completes the proof of the Isoperimetric Theorem by showing that a maximum figure must have such properties and therefore a center point and finally that all diameters are equal. His proof is open to the same objection as that of Steiner.

Padoa (1873-) has given an excellent review [40] of the first two proofs of Steiner.

Caratheodory (1873-) and Study (1862-1930) have written a joint paper [41] in which each gives a separate proof. They modify the method of Steiner so as to make a direct instead of an indirect proof. The change of method is made in order to avoid the necessity of making the assumption that a maximum exists.

An outline of the proof of Caratheodory is as follows: The plane is divided into halves by a line g, and from a point A on this line an arbitrary curve of length π is constructed

so that its end point is also on g. A modification of the method of Steiner is used to construct an infinite series of auxiliary curves all of length π and having a common beginning point A, but whose end points are a series of points on g. Since these end points are at most a distance π from A, it is possible to select a subset which has one and only one limit point ω. The areas inclosed between the curves and the line g are denoted by

$$I_1, \quad I_2, \quad I_3, \text{-------} \quad I_n, \text{----}$$

and the method of construction shows that no one is smaller in area than the one preceding. Since the areas of these curves have an upper bound and are increasing they must converge to a limit, and Caratheodory proves that these curves have as a limiting curve the semicircle drawn through A and ω. Since the lengths of the curves remain equal to π, the circle must have a radius equal to one. Finally if I_0 denotes the area inclosed between the first arbitrarily selected curve and the line g, then

$$I_0 = I_n = \pi/2 ,$$

and the equality sign can hold only when I_0 is a semicircle, which proves that a semicircle of length π incloses a larger area with g than any other curve of equal length.

Study makes use of the symmetric method of Steiner. He starts with a polygon which has a perimeter 2π in length and shows how to construct a new convex polygon that has n

axes of symmetry. For the new polygon the perimeter remains constant and the area is not less than the original one. As n increases there is produced a series of polygons that approach as a limit a circle whose radius r is to be determined. Since the perimeters of the polygons remain equal to 2π and are either equal to or greater than the perimeter of the limiting circle, it follows that $r \leqq 1$. Furthermore the areas of the polygons must approach the area πr^2, and hence the area of the originally given polygon must be $\leqq \pi$. If the equality sign holds, all the polygons of the series must be equal in area. But in that case it is pointed out that the method of construction leads to a contradiction. It follows that a circle of perimeter 2π has a larger area than any polygon of equal perimeter.

E. De Leber has given three proofs [42] based again on the postulate that there is a maximum figure among all those figures having a given perimeter.

Bieberbach (1886-) has given a proof of the theorem [44] that among all domains that are plane, finite, and have a given diameter the circle has the largest area, where he means by diameter the maximum distance between any two points of the border.

Minkowski (1864-1909) developed a theory [36] regarding pairs of parallel ovals that reduces to the Isoperimetric Theorem for a special case. This work is the beginning of

the second group of proofs mentioned in the first paragraph of this section. He defines an oval by means of a function $H(u,v)$ which appears in the inequality,

$$ux + vy \leq H(u,v).$$

The totality of such inequalities for all possible values of u and v defines a domain of points x,y which is called an oval. The area of an oval is

$$1/2 \int_0^{2\pi} H\left(H + \frac{d^2 H}{d\theta}\right) d\theta$$

where H is defined as follows: Draw a unit circle about the center of gravity of the oval and consider this center of gravity as the origin of the system of coordinates. For some point α, β on the perimeter of the unit circle write $\alpha = \cos\theta$, $\beta = \sin\theta$. Choose the point on the boundary of the oval where the outer normal has the direction α, β and write

$$H(\alpha,\beta) = H(\cos\theta, \sin\theta) = H.$$

Suppose that we have two such ovals defined by H_1 and H_2 and construct a third such that $H = (1-t)H_1 + tH_2$. If we denote the area of this third oval by F then

$$F = (1-t)^2 F_1 + 2(1-t)tM + t^2 F_2$$

where F_1 and F_2 are the areas of the first two ovals and M is given by the equation,

$$M = 1/2 \int_0^{2\pi} H_1\left(H_2 + \frac{d^2 H_2}{d\theta^2}\right) d\theta = 1/2 \int_0^{2\pi} H_2\left(H_1 + \frac{d^2 H_1}{d\theta^2}\right) d\theta.$$

This quantity M is an invariant with respect to all parallel translations. Minkowski calls M the mixed area of the two curves. It is proved that

$$(1) \quad M^2 - F_1 F_2 \geqq 0$$

and that the equality sign is valid only when the two ovals are similar (homothetic). If the second oval is chosen as a unit circle, H_2 becomes the integer one and the expression for M shows that 2M is the perimeter of the first oval. If the perimeter and area of the first oval are denoted by L and F then (1) gives

$$L^2 - 4\pi F \geqq 0$$

and the equality sign is valid only when the first oval is a circle. It is easily verified that this last sentence is a statement of the isoperimetric property of the circle.

The theory of parallel curves has been discussed by several writers. Crone (1877-) has an article [37] on this subject that appeared about one year after the work of Minkowski. Blaschke (1885-) gave a discussion [45] of the theory of Minkowski in an article that appeared in 1914. He discussed first the case when two quadrilaterals take the place of the ovals of Minkowski. The theory is then extended to polygons and finally to any closed curves. A paper [46] written by Frobenius (1849-1917) was published in 1915 and another [49] by Liebmann (1874-) in 1919. The last two papers use the theory of quadratic forms to prove the isoperimetric property. If the area F(t) of any curve of a family is expressed in terms of the area F and the perimeter L of an arbitrary curve, for example by a formula

$$F(t) = F + tL + \pi t^2,$$

then F(t) can be made negative for real values of t only when the discriminant of the quadratic form is greater than zero. The motive is then to show that for any closed curve that is not a circle a value of t exists for which F(t) is negative and therefore the discriminant is greater than zero. The greater part of Liebmann's paper is concerned with proving this last statement. Bonnesen (1873-) has given an account of the theory of the mixed area of two ovals in his book [53] published in 1929. Blaschke gave a proof [55] of the Isoperimetric Theorem in his Differential Geometry. His proof is similar to that of Crone and Frobenius, and is notable for its simplicity. It depends on a theorem as follows:

Theorem. There exists between the area F and the perimeter L of a circle the relation

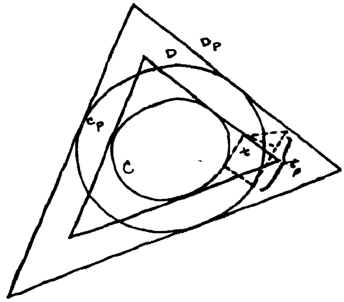
$$L^2 - 4\pi F = 0$$

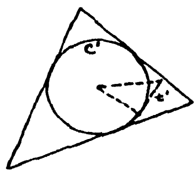
and for every other plane curve the relation is

$$L^2 - 4\pi F > 0.$$

Let C be an oval with the clockwise direction prescribed on it, and let C_p be the outer parallel curve at a distance p with the tangents to C_p parallel in the same sense to the tangents of C. C_p is evidently an oval. Let C' be a unit circle. Let a triangle of tangents to C be denoted by D whose area contains C. Let D_p and D' be the corresponding

triangles of parallel tangents for C_p and C'. Let t be the length of a line measured in a positive sense along a tangent from a point of contact on C to the intersection of this tangent with a side of D. Let t_p and t' be the corresponding lengths for parallel tangents to C_p and C'. Let ϕ be the angle between the tangent and a fixed direction. Let d, d_p, and d' be the areas of D, D_p, and D'. Let F, F_p, and F' be the areas of C, C_p, and C'.





Let r be the radius of the inscribed circle to the triangle D. Then we have the following relations,

$$t_p(\phi) = t(\phi) + pt'(\phi) \;,\; F_p = d_p - 1/2\int_{-\pi}^{\pi} t_p^2 \, d\phi \;,$$

$$(1) \quad F_p = (p + r)^2 d' - 1/2\int_{-\pi}^{\pi} (t + p\,t')^2 \, d\phi .$$

The area of the parallel curve may be expressed by the equation

$$(2) \quad F_p = F + pL + p^2 F'$$

where L denotes the perimeter of C. This can be seen as follows: Consider the small trapezoid inclosed by the arc length ds of C, the two lengths measured along the normals to C through the ends of ds, and the arc of the outer curve inter-

cepted by these lengths. The area of each of these small figures is

$$p(ds + 1/2\ p\ d\phi)$$

and therefore by taking the sum of them one sees that

$$F_p - F = \int p(ds + 1/2\ p\ d\phi) = pL + p^2\pi$$

For $p = 0$, F_p is positive by equation (2) and for $P + r = 0$ is less than or equal to zero by equation (1). Therefore the roots of $F_p = 0$ are real and it follows that

$$L^2 - 4\pi F \geq 0$$

If the equality sign holds F_p must not change in sign. It is positive for $p = o$, therefore if one sets $p = -r$ in (1) the integral must vanish and the ratio t/t' equal to r is constant. Therefore one can vary one side of the triangle of tangents and r will remain unchanged. It follows that all circumscribed triangles to the oval C have an inscribed circle such that the diameters of all the circles are equal. A figure such that all its diameters are equal must be a circle. The final remark of the author is that the proof is valid only when the perimeter of the figure contains no corners or straight lines.

Lebesgue (1875-) wrote a paper [43] in 1914 in which he shows that the Isoperimetric Problem may be stated in a new way. His statement of the problem is as follows: Find a domain for which the ratio L^2/F, the square of the perimeter to the area, shall be the smallest possible.

Blaschke has given a proof of the Isoperimetric Theorem by means of a method which is different from that described above. He has written two articles in which the same method is used. His paper [45] of 1915 is a brief account of a proof which is given more in detail in his book [47] which was published in 1916. It is proved that if one has a closed, continuous, rectifiable, plane curve of length L and area F one can construct a polygon such that if λ is its perimeter and ϕ its area then it is true that

$$|L-\lambda| < \epsilon, \quad |F-\phi| < \epsilon$$

where ϵ is an arbitrary quantity, and it also follows that

$$\lambda^2 - 4\pi\phi > 0$$

Now suppose that

$$L^2 - 4\pi F < 0$$

then one can construct a polygon such that

$$\lambda^2 - 4\pi\phi < 0$$

which is contrary to the results of a previous proof and therefore one must have

$$L^2 - 4\pi F \geqq 0$$

The method of Steiner may now be used to show that the equality sign holds only for a circle.

Bonnesen has written a number of articles [50,51,52] on the Isoperimetric Problem but he has given the essentials of all this material in a book [53]. He begins the discussion with the remark that it is sufficient to prove the Isoperimetric

Theorem for convex curves because it is easy to see that a curve which is not convex can not furnish a maximum. He calls the expression $L^2/4\pi - F$ the "Isoperimetric Deficit" of a curve. He states that the formulas

$$L_1 = L + 2\pi t \quad , \quad F_1 = F + Lt + \pi t^2$$

can be easily derived for the case of two parallel polygons, and for the general case it is sufficient to pass to the limit. He proves that

$$L^2/4\pi - F \geqq (\pi/4)(R-r)^2$$

for a convex curve, where R and r are the radii of the circumscribed and inscribed circles of the curve. Bernstein (1878-) in 1905 worked out a number of inequalities for curves on a sphere. One of the relations which he obtained for the case of a plane is the inequality of Bonnesen which has been written above except that the constant factor $\pi/4$ is replaced by $1/8(1 + 2\pi)^2$. Bonnesen refers to this inequality of Bernstein and says that it was the first inequality of this type to be derived.

A brief outline of the proof of Bonnesen is as follows: Consider the quadratic form

$$F + Lt + \pi t^2.$$

Its discriminant is precisely the deficit of the curve when one divides out the factor 4π. The deficit is greater than or equal to zero when the roots of the quadratic form are real. The form is positive for $t = 0$. If there exist values of t

for which the form is less than or equal to zero then the roots are real. If there exists a value of t for which

$$Lt - F - \pi t^2 \geqq 0,$$

then the negative of this value of t will make

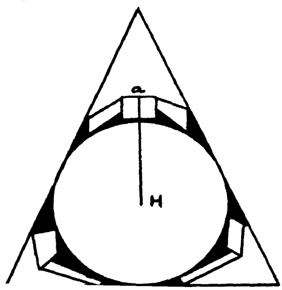
$$F + Lt + \pi t^2 \leqq 0.$$

The proof is complete when it is proved that

$$Lt - F - \pi t^2 \geqq 0$$

for t equal to r or R, the radii of the inscribed and circumscribed circles to the convex figure.

Bonnesen proves the last statement in the preceding paragraph when the figure is a convex polygon as follows: Inscribe a circle in the polygon touching two sides, and if the sides are parallel this is the largest circle that can be inscribed, but if not one can always draw the circle to touch three sides. Draw tangents to the circle at the three points where it touches the polygon. These tangents will form a triangle. Move each side of the polygon parallel to a fixed direction, say parallel to the bisectors of the angles of the triangle or parallel to any two parallel sides of the polygon until each side of the polygon is tangent to the circle. If a is the length of a side of the polygon, H its distance from the center of the circle,



and r the radius of the circle then one has

$$\Sigma\left[1/2\,Ha - (H - r)a\right] = (\Sigma\, ra - \Sigma\, 1/2\,Ha) \geqq \pi r^2$$

But

$$F = \Sigma\, 1/2\,Ha, \quad L = \Sigma\, a$$

and therefore

$$(1) \quad rL - F \geqq \pi r^2$$

and the equality sign is valid only when the figure is a circle. But (1) can be written in the form

$$(2) \quad L^2/4\pi - F \geqq \pi(L/2\pi - r)^2$$

When a similar process is carried out for a circumscribed circle another inequality is produced, namely,

$$(3) \quad L^2/4\pi - F \geqq \pi(R - L/2\pi)^2$$

The inequalities (2) and (3) may be combined into the inequality

$$(4) \quad L^2/4\pi - F \geqq \pi/4(R - r)^2$$

The Isoperimetric Theorem can be easily deduced from this result.

Bonnesen says that the method can be applied to any convex curve by moving all tangents and all straight lines of the perimeter parallel to a fixed direction as in the case of the polygon. He has given a second proof [53] by means of a symmetric method. He has also modified the symmetric method of Steiner. In each case he arrives at the result (4) given in the preceding paragraph.

3. Proofs by means of Fourier Series. The proof of the Isoperimetric Theorem by means of Fourier Series depends on a theorem which Hurwitz (1859-1919) called the fundamental theorem. This theorem is as follows: If

$$F(u) \sim a_0/2 + \sum_{1}^{\infty}(a_k \operatorname{Cos} k u + a'_k \sin k u)$$

and

$$H(u) \sim b_0/2 + \sum_{1}^{\infty}(b_k \operatorname{Cos} k u + b'_k \sin k u)$$

then

$$1/\pi \int_0^{2\pi} F(u)H(u)d u = a_0 b_0/2 + \sum_{1}^{\infty}(a_k b_k + a'_k b'_k)$$

where the sign $\sim$ is to be read equivalent to, and means that the right hand side of the first two equivalences represent the usual trigonometric development of the functions except that nothing is implied as to the convergence of the series or its equality with the function on the left. The equivalent sign may be replaced by the equality sign when the function defining the series has suitable continuity properties.

Let $x = x(s)$, $y = y(s)$, $0 \leqq s \leqq L$ be the equations of a simply closed, continuous, rectifiable curve where s is the length of arc and L is the length of the perimeter. It is known that a function which is continuous, periodic, and has a derivative which is continuous except at a finite number of finite discontinuities, may be expressed by a Fourier Series which converges uniformly and absolutely. Let us suppose that the functions for x and y which define the curve possess these properties. Then if we substitute for s the parameter u

given by the equation $u = 2\pi S/L$. x and y may be expressed as follows:

$$x = a_0/2 + \sum_1^{\infty}(a_k \operatorname{Cos} k u + a'_k \sin k u),$$

$$y = b_0/2 + \sum_1^{\infty}(b_k \operatorname{Cos} k u + b'_k \sin k u).$$

The derivatives with respect to u are:

$$dx/du \sim \sum_1^{\infty} k(a'_k \operatorname{Cos} k u - a_k \sin k u),$$

$$dy/du \sim \sum_1^{\infty} k(b'_k \operatorname{Cos} k u - b_k \sin k u).$$

When the variable s in the equation $(dx/ds)^2 + (dy/ds)^2 = 1$ is replaced by u the equation becomes

$$(dx/du)^2 + (dy/du)^2 = (L/2\pi)^2.$$

The integration of this equation gives

$$\int_0^{2\pi}\left[(dx/du)^2 + (dy/du)^2\right]du = 2\pi(L/2\pi)^2 = L^2/2\pi.$$

The application of the fundamental theorem described above to this last equation gives

$$(1) \qquad \pi \sum_1^{\infty} k\,(a_k^2 + a_k'^2 + b_k^2 + b_k'^2) = L^2/2\pi$$

The area inclosed by the curve may be represented by the equation

$$F = \int_0^{2\pi}(x\,dy/du)\,du.$$

But again on account of the fundamental theorem this reduces to

$$(2) \quad F = \pi \sum_1^{\infty} k(a_k b'_k - a'_k b_k).$$

It follows from (1) and (2) that

$$L^2/2\pi - 2F = \pi \sum_1^{\infty}\left[(ka_k - b'_k)^2 + (ka'_k + b_k)^2 + (k^2-1)(b_k^2 + b_k'^2)\right],$$

and since the right hand side of this equation is never negative it must be true that

$$L^2 - 4\pi F \geq 0$$

One observes that the equality sign is valid only when

$a'_1 + b_1 = 0,\ a_1 - b'_1 = 0,\ a_k = a'_k = b_k = b'_k = 0 \quad (k = 2,3,4----).$

But then the series for x and y reduce to

$$x = a_0/2 + a_1 \operatorname{Cos} u + a'_1 \sin u,$$

$$y = b_0/2 - a_1 \operatorname{Cos} u + a_1 \sin u$$

which are the parametric equations of a circle. Therefore for all simply closed, rectifiable, plane curves whose equations $x = x(s)$, $y = y(s)$ have the properties described at the beginning of this paragraph it is true that

$$L^2 - 4\pi F \geqq 0$$

and the equality sign is valid only when the curve is a circle. The Isoperimetric Theorem follows easily from this inequality.

The first proof [56] of the Isoperimetric Theorem which made use of the Fourier Series was given by Hurwitz. Lebesgue has also given a proof [57] similar to that given by Hurwitz. Each of these proofs is preceded by a detailed discussion of the theory of the Fourier Series. The application of the theory to the proof of the Isoperimetric Theorem is in each case the same except for a few details. The type of curves to which the proofs apply is determined by the development of the theory of the Fourier Series. Hirschler has written a thesis [58] in which the area of a regular closed curve is compared with that of a circle by means of the Fourier Series. He defines a regular curve as follows: A regular curve is continuous, consists of a finite number of arcs which do not

intersect themselves, and each arc possesses a continuously turning tangent at every interior point and end point. He applies the fundamental theorem which is stated in the first paragraph of this section to the symmetric formula for the area

$$A = 1/2\int_0^{2\pi}(xy' - yx')\, d\, S,$$

and uses the integral of the relation

$$(dx/ds)^2 + (dy/ds)^2 = 1$$

to compute the area of a circle which has a perimeter equal to the perimeter of the regular curve in terms of the Fourier Constants of $x(s)$ and $y(s)$. He proves that the difference between the area of a circle of given perimeter and the area of a regular curve which has an equal perimeter is a positive quantity unless the regular curve is itself a circle. The proof is again preceded by a development of the Fourier Series. A good outline [55] of the proof of Hurwitz is given by Blaschke in his Vorlesungen über Differential Geometrie.

4. Proofs by means of the calculus of variations. A statement was made in Section 1 that the early proofs of the Isoperimetric Theorem which applied the calculus of variations were concerned only with the necessary condition for a maximum. This statement continues to be true for all such proofs which were published before the work of Weierstrass (1815-1897) was known. Many of the writers discussed a modification of our problem which we will call Problem I, and of which our problem

is a special case.

Problem I. To find among all curves which have a given length and join two fixed points on a straight line that one which incloses with the line a maximum area.

A first necessary condition that a simply closed, regular curve of given length enclose a maximum area, if such a curve exists, may be derived from the theory of Problem I by assuming that the two points on the line coincide. A first necessary condition on a solution of Problem I is that the maximizing curve shall satisfy the first necessary conditions that the integral

$$(1) \qquad \int_{t_1}^{t_2} \left[1/2(xy' - yx') + \lambda \sqrt{x'^2 + y'^2} \right] d\,t$$

have a maximum value where λ is a suitably selected constant. If we write H for the integrand of the integral, these conditions are that the derivatives $H_{x'}$, $H_{y'}$, be continuous on the maximizing arc and that the differential equations of Euler must be satisfied. The conditions that $H_{x'}$, and $H_{y'}$, are continuous imply that the maximizing arc has no corners. The differential equations of Euler are

$$(2) \qquad \begin{aligned} H_x - \tfrac{d}{dt} H_{x'} &= y' - \lambda \tfrac{d}{dt} \left(x'/\sqrt{x'^2 + y'^2} \right) = 0, \\ H_y - \tfrac{d}{dt} H_{y'} &= -x' - \lambda \tfrac{d}{dt} \left(y'/\sqrt{x'^2 + y'^2} \right) = 0. \end{aligned}$$

When they are integrated we find

$$y - b = \lambda y'/\sqrt{x'^2 + y'^2} \,; \quad x - a = -\lambda x'/\sqrt{x'^2 + y'^2} \,;$$
$$(x-a)^2 + (y-b)^2 = \lambda^2 .$$

Therefore the solution must be an arc of a circle. It is clear

that the method is the same when the two points on the line coincide.

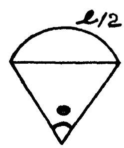
Proofs arriving essentially at the result above have been given by Bordoni (1789-1860) [59], Dienger (1818-1894) [63], Jellett (1817-1888) [60], Moigno-Lindelof [61], and Lundstrom [62], and more recently by Thomé (1841-1910) [68], and Hadamard (1865-) [70].

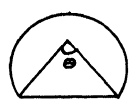
Weierstrass was the first to make a complete proof by the method of the calculus of variations that the area inclosed by a circle is larger than that inclosed by any other regular closed curve of equal length. His proof also precedes the other complete proofs by other methods. Unfortunately he did not publish his work. The record of it was preserved however, by means of the lecture notes of his students. The seventh volume of his collected works [74] based on several collections of these notes, was published in 1927.

Schwarz (1843-1921) used a method [64] similar to the method of Weierstrass to show that among all closed curves which inclose a given area the circle has the shortest perimeter. It was pointed out in Section 2 that the Isoperimetric Theorem follows from this result.

Kneser (1862-1930) has a discussion of the Isoperimetric Theorem in his book [65] on the calculus of variations. He develops both the necessary and the sufficient conditions for a maximum for Problem I. In the second edition of his book

he deduces the Isoperimetric Theorem from the solution for Problem I. His argument is as follows: Consider a closed curve consisting of a finite number of regular parts, and let the length of the curve be l. Find two points P_1, P_2 which divide the perimeter into two equal parts and join these points with a straight line. Because of the solution to Problem I, the points may be joined by arcs of circles each of length l/2 in a manner such that the area inclosed by the two arcs is greater than the area inclosed by the original curve. It remains to be proved that a circle which has a circumference equal to l incloses a larger area than the area inclosed by the two arcs of circles the sum of whose lengths is l, and which join the two points P_1, P_2. Kneser does this by proving that a semicircle which has a length l/2 incloses a larger area with its chord than any other arc of length l/2 incloses with the same chord. To prove this consider an arbitrary arc of a circle which has length l/2 and join the ends of this arc by a line. Draw the radii of the arc at the two ends. Call the central angle between these two radii θ and denote the area between the arc and its chord by F. Then the area F





may be expressed by the equation

$$F = \theta/2(l/\theta)^2 - (l/\theta)^2 \sin\theta/2 \cos\theta/2 = \left(\frac{\theta - \sin\theta}{\theta^2}\right).$$

Take the derivative of F with respect to θ, set this derivative equal to zero, and divide out the constant factor $l^2/8$. This gives

$$\frac{-\theta^2 - \theta^2\cos\theta + 2\theta\sin\theta}{\theta^4} = \frac{-2\theta^2\cos^2\theta/2 + 4\theta\sin\theta/2\cos\theta/2}{\theta^4}$$

$$= 4/\theta \quad \cos\theta/2(\sin\theta/2 - \theta/2\cos\theta/2) = 0.$$

This has only one solution π between 0 and 2π and this solution gives a maximum. But when the central angle is equal to π the figure is a semicircle. Therefore a semicircle incloses with its chord a larger area than any other arc of equal length. It follows that a circle constructed from two such semicircles possesses a perimeter equal to the sum of the lengths of the two arcs joining the two points on the original curve and an area greater than the area included between the two arcs.

Hancock (1867-) gave a discussion of the theory of Weierstrass in a book on the calculus of variations which was published in 1904. He gave a proof [67] of both the necessary and sufficient conditions for a closed curve to inclose a maximum area.

Bolza (1857-) has given an excellent discussion [69] of Problem I. His derivation of the first necessary condition is different from that given in the third paragraph of this

section. He reduces the differential equations of Euler to the equivalent form

$$H_x y' - H_y x' + H_1(x' y'' - y' x'') = 0,$$

where

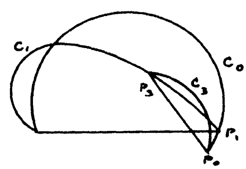
$$H_1 = \lambda/(\sqrt{x'^2 + y'^2})^3$$

and therefore the differential equation reduces to

$$1/r = (x' y'' - y' x'')/(\sqrt{x'^2 + y'^2})^3 = -1/\lambda .$$

This shows that λ is different from zero and that the curve sought must be an arc of a circle.

The sufficiency proof of Bolza for Problem I is as follows: Draw a circular arc C_0 of length 21 through the points P_1, P_2. Near this draw an arbitrary, admissible curve C' of the same length through the same two points. Let P_0 be a point on the extension of C_0 and suppose that it is not a point of C'. Choose an arbitrary point P_3 on C'. Then P_3 is different from P_0 and the sum Z_3 of the length of the arcs P_0P_1 plus P_1P_3 is larger than the distance P_0P_3. There can be drawn only one arc of a circle C_3 through P_0 and P_3 which has a length equal to Z_3 and which is described in a positive sense. Bolza has shown in a previous paragraph that a congruence of space extremals through P_0 may be defined by the equations



$$x - x_0 = -2\lambda \cos(T+K)\sin T, \quad y - y_0 = -2\lambda \sin(T+K)\sin T, \quad Z = -2\lambda T,$$

where T, λ, K are limited to the domain,

$$0 < T < \pi, \quad 0 \leqq K < 2\pi, \quad \lambda < 0.$$

This congruence makes a space field S which on account of the inequality

$$z_3 > \sqrt{(x - x_0)^2 + (y - y_0)^2} > 0$$

fills up a definite part of space and C' lies wholly in this field. The E-Function of Weierstrass for this problem is

$$\mathcal{E}(x_3, y_3; p_3, q_3; p_3, q_3; \lambda_3) = \lambda_3[1 - \cos(\theta_3' - \theta_3)]$$

where λ_3 is negative and $\cos(\theta'_3 - \theta_3)$ is not zero along the whole curve C', since we suppose that C' is different from C_0. The points conjugate to P_0 may be expressed by the zero values of a certain determinant. This determinant for this problem may be reduced to

$$\Delta(T_3, K_3, \lambda_3) = 8\lambda_3^2 \sin T_3(\sin T_3 - T_3 \cos T_3).$$

It is clear that the determinant is not equal to zero when

$$0 < T_3 < \pi, \quad \lambda_3 < 0.$$

Therefore it follows from the theory of Weierstrass that the integral (I) along the arc of the circle C_0 is greater than the same integral along any other curve C' which is admissible and different from C_0 This says that the solution to Problem I is an arc of a circle.

Bolza does not extend the method so that it applies to a complete circle but states that such an extension can be made.

After one has found the conditions necessary that a curve inclose a larger area than any other curve of equal length the proof may be made complete by proving that these conditions, or a modification of them, are sufficient to insure

a maximum, or one may use a second method which consists in proving that there exists a curve of given length which incloses a maximum area. Tonelli (1885-) has used the second method [72] in proving the Isoperimetric Theorem.

Bonnesen devotes a chapter in his book [75] to the proof of the Isoperimetric Theorem. He uses polar tangential coordinates and the method of Weierstrass. He states his results as follows: Let E be a circle of radius r, (T) the complete ensemble of triangles and pairs of parallel lines circumscribed to E. Let C be the ensemble of convex curves which may be inscribed in one of the figures T. Between the perimeter L and the area F of any one of these curves C one has the inequality

$$rL - F \geqq \pi r^2$$

and the equality sign is valid only when C is a circle. This agrees with his result which has been described in Section 2.

Another problem closely related to the Isoperimetric Theorem is the problem which has been called the problem of Dido. It may be stated as follows: To find among all curves of given length which join two points on an arbitrary curve one which incloses with the arbitrary curve a maximum area. If the arbitrary curve is a straight line and the points are fixed it reduces to Problem I, but if the end points are variable it reduces to a new problem. It is known that if the end points are variable the maximum curve must intersect the arbitrary

curve at right angles. If the arbitrary curve is a straight line the maximum curve is a semicircle. It is easy to deduce the Isoperimetric Theorem from this result. Kneser has made an extensive study [66] of the problem for the case of variable end points. Merrill (1887-) has also written a paper [71] on the variable end point case. There is also a paper by G. Weyl [73] in which sufficient conditions for a more general problem of the calculus of variations are applied to the problem of Dido.

BIBLIOGRAPHY

GENERAL REFERENCES

1. Lecat, Encyclopédie des Sciences Mathématiques, II (1913), pp. 232-243.

2. Kneser, Encyklopädie der Mathematischen Wissenschaften, II A 8 (1900), pp. 608-609, 611-616; Zermelo und Hahn, ibid., II A 8a (1904), pp. 635-638.

3. Cantor, Vorlesungen über Geshichte der Mathematik, vol. 1 (1892), pp. 167, 341-342, 418, 696; vol. 2, p. 105.

4. Montucla, Histoire des Mathématiques, vol. 1 (1799), p. 117.

5. Allman, Greek Geometers from Thales to Euclid (1889), p. 46.

6. Libri, Histoire des Sciences Mathématiques En Italie, vol. 3 (1865), pp. 232-234.

7. Heath, A History of Greek Mathematics, vol. II (1921), pp. 205-213.

8. Chisini, Sulla Teoria Elementare Degli Isoperimetri, in Enriques, Questioni Riguardanti Le Mathematiche Elementari, vol. 3, pp. 201-310.

1. ORIGIN OF THE PROBLEM AND EARLY PROOFS

9. Aristotle, De Caelo, p. 287[a]; The Works of Aristotle translated into English by J. A. Smith and W. D. Ross, vol. 2, part 2 (1908), p. 287[a].

10. Polybius, The Histories, book IX, p. 26[a]; English translation by W. R. Paton, vol. IV (1922-1927), p. 61.

11. Diogenes Laertius, Lives of Eminent Philosophers, English translation by Yonge (1901), p. 352.

12. Proclus, Lycius, surnamed Diadochus, The Commentaries of Proclus on the Timaeus of Plato, p. 384; English translation by Taylor (1820), p. 450.

13. Proclus, Lycius, surnamed Diadochus, Commentary on Book I of Euclid, p. 403; See also reference 7. pp. 206-207.

14. Simplicius, Commentaries on the De Caelo of Aristotle, Edition of Heiberg, vol. II (1894), p. 412; A quotation in Greek with a translation in German in reference 31, p. 5.

15. Theon of Alexandria, Θέωνος Ἀλεξανδρέως εἰς τά αὐτα ὑπομνημάτων βιβλία ια, Bale 1538.

16. Pappus of Alexandria, Pappi Alexandrini Collectiones, Pise (1588), Venise (1589), Book 5, pp. 83-115; Edition of Hultsch (1878), vol. 1, pp. 323-337, vol. 3, pp. 1189-1194.

17. Christopher Clavius, In Sphaeram Joannis De Sacro-Bosco Commentarivs, Venise (1591), pp. 97-98; 3'd Edition Saint-Gervais (1608), pp. 110-113.

18. Kepler, J., Nova Stereometria Doliorum Venariorum, Linz (1615), part II, theorem IV; Opera, Edition of C. Frisch (1863), p. 607.

19. Galileo Galilei, Two New Sciences, English translation by Henry Crew and Alfonso De Salvio (1914), pp. 58-60.

20. Barrow, I., Archimedes Opera Methodo Nova Illustrata (1675), p. 262.

21. Bernoulli, James, Opera, vol. 2 (1744), p. 775.

22. Euler, Leonhard, Commentarii Academiae Scientiarum Imperialis Petropolitanae, vol. 6 (1732-1733), Edition 1738, pp. 142-143.

23. Elvius, P., K. Vetensk. Akad. Handl. Stockholm, vol. 2 (1741), Edition 1743, pp. 136-137; German translation by A. G. Kastner, Schwed. Akad. Wiss. Abh., vol. 3 (1741), Edition 1750, pp. 160-161.

25. Simpson, T., An Investigation of a General Rule for the Resolution of Isoperimetrical Problems of all Orders, Philosophical Transactions, London, vol. 49 (1755), Edition 1756, pp. 4-15; vol. 50 (1758), Edition 1759, pp. 623-631.

26. Riccati, G., Delle figure piane isoperimetre contenenti le massime superficie, dissertazione, Nuova Raccalta D'Opuscoli scientifici e filologici, vol. 18 (1772), pp. 2-14.

27. L'Huiler, S., De problematibus quae vocantur isoperimetricis, Principiorum Calculi Differentialis et Integralis (1795), pp. 321-322.

28. Woodhouse, Robert, A Treatise on Isoperimetric Problems and the Calculus of Variations (1810), pp. 41-42.

29. Simpson, T., Elements of Geometry (1821), pp. 195-199.

30. Legendre, Elements of Geometry, English translation by John Farrar (1885), pp. 99-104.

31. Schmidt, Wilhelm, Sur Geschichte der Isoperimetrie im Altertume, Bibliotheca Mathematica, vol. 22 (1901), pp. 5-8.

2. GEOMETRIC PROOFS OF THE ISOPERIMETRIC PROPERTY OF THE CIRCLE

32. Anonymous, Annales De Mathématiques Pures ET Appliques, vol. 13 (1822-1823), pp. 132-139; vol. 14 (1823-1824), pp. 316-317.

33. Steiner, J., Einfache Beweise der isoperimetrischen Hauptsatze, Crelle's Journal, vol. 18 (1838), pp. 286-287; Sur le maximum et minimum des figures dans le plan, sur le sphere, et l'espace en general, ibid., vol. 24 (1842), pp. 105-106, 197, 205, 209.

34. Edler, F., Vervollstandigung der Steiner'schen elementar-geometrischen Beweise für den Satz, dass der Kreis grosseren Flachenhalt besitz als jede andere Ebene Figur gleichgrossen Umfanges, Göttingen Nachrichten (1882), pp. 73-80; Translated into French and printed in the Bulletin Des Sciences Mathématiques, second series, vol. 7 (1883), pp. 198-204.

35. Sturm, R., Bemerkungen und Zusatze zu Steiner's Aufsatzen über Maximum und Minimum, Crelle's Journal, vol. 96 (1884), pp. 50-53; Maxima und Minima in der elementaren Geometrie (1910), pp. 31-51.

36. Minkowski, H., Volumen und Oberflache, Mathematische Annalen, vol. 57 (1903), pp. 459-467.

37. Crone, C., Om Prismatoidens Rumfang, Nyt. Tidsskrift für Mathematik, 15 B (1904), pp. 73-75.

38. Bernstein, Felix, Über die isoperimetrische Eigenschaft des Kreises, Mathematische Annalen, vol. 60 (1905), pp. 117-136.

39. Witting, A., Ein Beitrag zum isoperimetrischen Problem in der Ebene, Archiv der Mathematik und Physik, series 3, vol. 12 (1907), pp. 288-290.

40. Padoa, A., Une question de maximum, Nouvelles Annales De Mathématiques, series 4, vol. 8 (1908), pp. 529-535.

41. Caratheodory, C. and Study, E., Zwei Beweise des Satzes, das der Kreis unter allen Figuren gleichen Umfanges den grossten Inhalt hat, Mathematische Annalen, vol. 68 (1910), pp. 133-140.

42. De Leber, Edler, Le contenu du cercle et le sphere, L'Enseignement Mathématiques, vol. 15 (1913), pp. 369-384.

43. Lebesgue, Sur les problemes des isopérimètres et sur les domaines de largeur constante, Comptes Rendus Des Séances (1914), pp. 72-76.

44. Bieberbach, Ludwig, Über eine Extremaleigenschaft des Kreises, Jahresbericht Der Deutschen Mathematiker-Vereinigung, vol. 24 (1915), pp. 247-250.

45. Blaschke, W., Beweise zu Satzen von Brunn und Minkowski über die Minimaleigenschaft des Kreises, Jahresbericht Der Deutschen Mathematiker-Vereinigung, vol. 23 (1914), pp. 210-234; Kreis und Kugel, ibid., vol. 24 (1915), pp. 195-199.

46. Frobenius, G., Über den gemischten Flachenhalt zweier Ovale, Sitzungsberichte, Berlin (1915), pp. 387-404.

47. Blaschke, W., Kreis und Kugel (1916), pp. 1-32, in particular pp. 30-31.

48. Lorey, W., Über isoperimetrische Probleme in der Schule und in der Forschung, Zeitschrift fur Mathematischen und Naturqissenschaftlichen Unterricht, vol. 49 (1918), pp. 281-293.

49. Liebmann, Heinrich, Das Frobeniussche Kappendreieck und die isoperimetrische Eigenschaft des Kreises, Mathematische Zeitschrift, vol. 4 (1919), pp. 288-294; Die isoperimetrische Eigenschaft des Kreises, Sitzungsberichte Math. Ph. Kl. Bayer Akademie, Munchen (1919), pp. 111-114.

50. Bonnesen, T., Sur une amelioration de l'inequality isopérimètrique du cercle et demonstration d'une inequalite de Minkowski, Comptes Rendus, vol. 172 (1921), pp. 1087-1089.

51. Bonnesen, T., Ueber eine Verscharfung der isoperimetrischen Ungleichheit des Kreises in der Ebene und auf der Kugeloberflache nebst einer Anwendung auf eine Minkowskische Ungleichheit fur konockse Korper, Mathematische Annalen, vol. 84 (1921), pp. 216-227.

52. Bonnesen, T., Ueber das isoperimetrische Deficit ebener Figuren, Mathematische Annalen, vol. 91 (1924), pp. 252-256; Cirklens isoperimetriske ulighed, Mat. Tidsskrift (1925), A, pp. 53-61.

53. Bonnesen, T., Les Problemes Des Isopérimètres Et Isepiphanes (1929), pp. 60-84.

55. Blaschke, W., Vorlesungen über Differential Geometrie (1930), pp. 55-61.

3. PROOFS BY MEANS OF FOURIER SERIES

56. Hurwitz, A., Sur quelques applications géomètriques des séries de Fourier, Anneles L'Ecole Normale Superieure, vol. 19 (1902), pp. 392-394.

57. Lebesgue, H., Leçons sur les séries trigonomètriques (1906), pp. 107-110.

58. Hirschler, E. J., An Application of the Fourier Series to a Special Isoperimetric Problem, Master's Thesis, The University of Chicago (1916), pp. 28-32.

4. PROOFS BY MEANS OF THE CALCULUS OF VARIATIONS

59. Bordoni, A. M., Sulle figure isoperimetre esistenti su qualservaglia superficie, Opuscoli matematici e fisici di diversi autori, vol. I (1832), p. 9.

60. Jellett, Calculus of Variations (1850), pp. 154-160.

61. Moigno-Lindelöf, Lecons sur le calcul differential et integral, 4, calcul des variations (1861), pp. 203-204, 233-235.

62. Lundström, C. E., Utkast til isoperimetriska Problemers fullstandiga Solution, Dissertation Upsala University (1866), pp. 33-36.

63. Dienger, J., Grundriss der Variationsrechnung (1867), pp. 52-54.

64. Schwarz, H. A., Ueber eine Eigenschaft der Kugel, Gesammelte Mathematische Abhandlungen, vol. 2 (1890), pp. 327-335; The same material in Göttingen Nachrichten (1884), pp. 1-13.

65. Kneser, A., Lehrbuch Der Variationsrechnung (1900), pp. 123-125, 136-140; Edition (1925), pp. 167-168, 179-181, 196-205.

66. Kneser, A., Die isoperimetrische Aufgabe wenn beide Endpunkt variabel sind, Mathematische Annalen, vol. 56 (1903), pp. 199-232.

67. Hancock, Harris, Lectures on the Calculus of Variations (1904), pp. 224-230, 288-291.

68. Thomé, L. W., Ueber eine Anwendung der Theorie der linearen Differential gleichengen in der Variationsrechnung, Journal für Mathematik, vol. 132 (1907), pp. 155-156.

69. Bolza, O., Vorlesungen über Variationsrechnung (1909), pp. 465-466, 483-484, 494-496, 498-499, 502, 510-511.

70. Hadamard, Leçons sur le calcul des variations (1910), pp. 202-204.

71. Merrill, A. S., An Isoperimetric Problem with Variable End Points, American Journal of Mathematics, vol. 41 (1919), pp. 60-78.

72. Tonelli, L., Fundamenti di calcalo delle variazioni, vol. II, pp. 542-543.

73. Weyl, G., Hinreichende Bedingungen des Extremums bei einer neuen Art von isoperimetrischen Problemen, Journal für Mathematik, vol. 152 (1923), pp. 76-79.

74. Weierstrass, Karl, Vorlesungen über Variationsrechnung (1927), pp. 257-264, 301-302.

75. Bonnesen, T., See Reference 53 above, pp. 155-166.